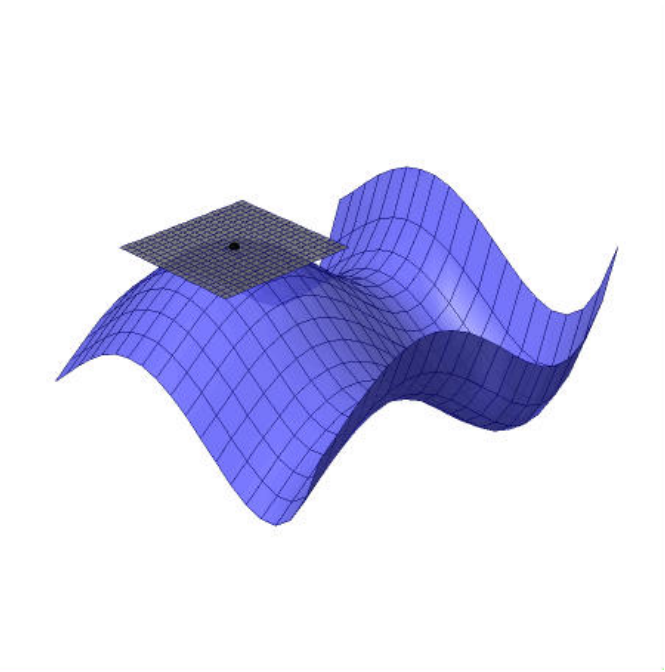


University College London

Autumn Term 2012

Mathematical Methods 3

Jonny Evans

Last updated January 15, 2013

Contents

Introduction and notation

This course is an introduction to some of the most ingenious ideas in 18th century mathematics. Some of these ideas are so important that their discoverers' names appear on the side of the Eiffel Tower.

If you perform sufficiently well on the final exam, who knows where your name will end up?

Structure

The course is divided into two parts.

In the first part we will see how to turn some natural physical and geometric problems into differential equations. Some of these will be ordinary differential equations (for functions of a single variable) others will be partial differential equations (for functions of several variables). How do these equations arise? In ordinary calculus of a single variable, maxima and minima of a function $\phi\colon \mathbf{R} \to \mathbf{R}$ at $x_0 \in \mathbf{R}$ are found by differentiating and finding the zeros of the

resulting function. The result is an equation for x_0:

$$\frac{d\phi}{dx}(x_0) = 0$$

Now suppose that you're interested in minimising or maximising something over a space of all functions! Suddenly you have to differentiate a *functional*, $F(\phi)$ i.e. a function which eats a function and outputs a number. The resulting *Euler-Lagrange* equation

$$\frac{\partial F}{\partial \phi}(\phi_0) = 0$$

is usually a differential equation for ϕ_0. If ϕ_0 itself is a function of several variables then it is likely that this Euler-Lagrange equation is a partial differential equation.

The physical and geometric problems we are interested in are very natural, for example:

- What is the shortest path between two points in the plane? (i.e. minimise the length functional over all paths).
- What is the shape in the plane with maximal area amongst shapes of a fixed perimeter? (i.e. maximise area bounded under the constraint of fixed perimeter).
- What is the function $\phi\colon U \to \mathbf{R}$ ($U \subset \mathbf{R}^2$) which minimises the surface area of its graph subject to the boundary condition that $\phi|_{\partial U}$ is some fixed function?

The first two give rise to ordinary differential equations (for example, the second becomes a harmonic oscillator equation for the (derivatives of the) components of the parametric curve bounding the shape). The last gives a quasilinear second order partial differential equation. I assume that you know how to solve harmonic oscillators and other ordinary differential equations, however...

...The second part of the course will develop techniques to solve some partial differential equations. We begin by treating first order equations (only involving first partial derivatives). These can be tackled by a technique called the method of characteristics which originates in geometric optics. In optics with the speed of light set equal to 1, there is a first order PDE called the eikonal equation

$$\left(\frac{\partial \phi}{\partial x}\right)^2 + \left(\frac{\partial \phi}{\partial y}\right)^2 = 1$$

satisfied by the function $\phi(x, y)$ whose value at $(x, y) \in \mathbf{R}^2$ is the length of time it take light to travel to (x, y) from some fixed light-emitting curve $C \subset \mathbf{R}^2$ (the light being emitted in a normal direction to C). The obvious way to solve this is to draw a straight line connecting C to (x, y) and meeting C at ninety degrees. Let t be the parameter for this line $(x(t), y(t))$ and observe that if we define $f(t) = \phi(x(t), y(t))$ then $f(t)$ satisfies the ODE $df/dt = 1$ (simply because it

takes light time t to travel t units because the speed of light is 1). I'm not saying that this is obvious from the eikonal equation, but from the physical interpretation of the solution ϕ it's actually a tautology.

The method of characteristics generalises this example and constructs, for any first order PDE, a system of "characteristic curves" (in this example the light rays emitted normally from C) along which the PDE reduces to an ODE.

Finally we move on to tackle second order equations. We will concentrate on three which arise very naturally in physics and geometry:

- The heat equation
$$\frac{\partial \phi}{\partial t} = \frac{\partial^2 \phi}{\partial x^2}.$$
- Laplace's equation
$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0.$$
- The wave equation
$$\frac{\partial^2 \phi}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2}.$$

These are linear and therefore highly amenable to solution. They also typify three different classes of second order equations with vastly different behaviours: parabolic, elliptic and hyperbolic.

The general strategy of this final part of the course will be the following:

- Find infinitely many special solutions to the equation. Maybe they don't satisfy the boundary conditions you're interested in, but don't worry. This step uses a clever idea called separation of variables: you assume the solution has the form $\phi(x, y) = X(x)Y(y)$, that is a product of two functions each depending on only one of the variables. Then the PDE reduces to a pair of ODEs which are easy to solve. Of course, separation of variables rarely works, but with these simple linear equations it is fiercely effective.
- Take infinite linear combinations of these special solutions to fit to your boundary conditions. By linearity of the PDE one can take linear combinations of solutions and obtain a function which still solves the equation. Fitting infinite sums of special functions (like sine and cosine) to arbitrary functions is the subject of Fourier theory and indeed this is how Fourier discovered Fourier theory.

There are also several appendices: the first is a recap of Fourier theory because we will use lots of Fourier series in the second half of the course; the second includes much of the Sage code I used to create the diagrams (I also used Inkscape); the other appendices comprise all the problem sheets for the course.

Videos

There are supplementary videos for this course, available on YouTube. These cover a mixture of basic material prerequisite for understanding the course and extension material for those who care to watch it. Links to these videos are in the relevant places in the notes.

What you should already know

You should be:

- Comfortable computing Fourier series and manipulating trigonometric expressions. If you're struggling, try working through the Schaum Outline Series volume on Fourier Analysis.
- Able to solve a wide variety of ordinary differential equations. At the very least the simple harmonic equation $X'' = \lambda X$! If you're struggling, try working through the Schaum Outline Series volume on Ordinary Differential Equations.
- Able to remember and apply Green's theorem and have a rudimentary knowledge of vector calculus (div, grad and curl, tangent vectors to curves etc.).
- Ready to try solving a problem for which you have not been given a precise template (most important).

Notation

> Sections, like this one, which are marked with a left bar contain material I consider to be nonexaminable, usually proofs. That doesn't mean you should ignore them, just that you shouldn't sit there memorising them getting your knickers into a twist.

Partial derivatives are written with curly dees $\frac{\partial f}{\partial x}$. Ordinary derivatives are written with Latin dees $\frac{df}{dt}$. The only difference is that $\frac{d}{dt}$ means the function being differentiated is a function of a single variable. People who write partial derivatives with Latin dees are wrong and will be penalised. People who write ordinary derivatives with curly dees on the basis that single variable is a special case of several variables are smart, but still wrong, because it's misleading for the reader. If ever I write d/dt instead of $\partial/\partial t$ and you don't know why I did it, **challenge me** (there's always the possibility that I made a mistake).

Also, if $x(t)$ is a function of time t, the notation $\dot{x}$ indicates differentiation with respect to t. If $y(x)$ is a function of x, the notation y' indicates differentiation

with respect to x. It's possible that $'$ might denote differentiation with respect to another variable like t. It shouldn't bother you because it will only ever occur when the function in question is a function of a single variable, so x' or $\dot{x}$ just mean "differentiate x with respect to whatever it depends on".

Acknowlegdements

This course owes a lot to the course taught in previous years by Dr R Bowles and Dr G Esler. Their notes have been indispensible in preparing these notes; many of the nice questions from their problem sets have made it onto these problem sets. Arnold's book "Lectures on Partial Differential Equations" provided most of the inspiration for my presentation of the method of characteristics; Spiegel's "Fourier Analysis" has been useful too: I learned Fourier series from it myself, long ago. Thanks most of all to my extremely diligent markers (Giancarlo Grasso, Abbygail Shaw and Huda Ramli) and my wonderful and inquisitive second years for pointing out so many errors and typos in the lectures, in the notes and on the question sheets.

Part I

Critical points in finite- and infinite-dimensional problems

Chapter 1

Calculus with several variables

For most of what I say, we'll take "several" to mean "two" to save on notation. There are no new ideas introduced by adding more variables. Henceforth we'll use (x, y) to denote Cartesian coordinates on the plane $\mathbf{R}^2$.

1.1 First derivatives

1.1.1 Partial derivatives

Given a function $f\colon \mathbf{R}^2 \to \mathbf{R}$, its partial derivatives are what we get by differentiating with respect to one of the variables x or y and keeping the other fixed. In other words, to get the partial derivative with respect to x, we restrict f to the horizontal lines $y = y_0$ (y_0 is some constant) and we get a function of one variable

$$f(x, y_0),$$

which we know how to differentiate (assuming it's differentiable!). We define the first partial derivative of f with respect to x at (x_0, y_0) as

$$\begin{aligned}\frac{\partial f}{\partial x}(x_0, y_0) &\equiv \frac{d}{dx}\bigg|_{x=x_0} f(x, y_0) \\ &\equiv \lim_{h\to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}.\end{aligned}$$

Similarly

$$\begin{aligned}\frac{\partial f}{\partial y}(x_0, y_0) &\equiv \left.\frac{d}{dy}\right|_{y=y_0} f(x_0, y) \\ &\equiv \lim_{k\to 0} \frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k}.\end{aligned}$$

Example 1. *1. If* $f(x, y) = x^2y^3$ *then* $\frac{\partial f}{\partial x} = 2xy^3$ *and* $\frac{\partial f}{\partial y} = 3x^2y^2$.

2. If $f(x, y) = \sin(x + y)$ *then* $\frac{\partial f}{\partial x} = \cos(x + y) = \frac{\partial f}{\partial y}$.

We will sometimes abbreviate $\frac{\partial f}{\partial x}$ to $\partial_x f$ or just f_x. Often you will see it written $f_{,x}$ to distinguish the subscript from an ordinary subscript (like a vector index).

1.1.2 Directional derivative

Why should we restrict attention to the lines $x = x_0$ or $y = y_0$? Let's pick a vector $v \in \mathbf{R}^2$ and look at all lines pointing in the direction v, i.e. all lines of the form

$$\{p + tv \ : \ t \in \mathbf{R}\}.$$

Suppose we want to compute the derivative of f at $p = (x_0, y_0)$ in the direction of $v = (v_1, v_2)$. Restrict f to the line $(x_0 + tv_1, y_0 + tv_2)$ and differentiate with respect to t to get what we call the directional derivative of f in the direction v at the point p:

$$v_p(f) \equiv \left.\frac{d}{dt}\right|_{t=0} f(x_0 + tv_1, y_0 + tv_2)$$

Certainly if $v = (1, 0)$ then we get

$$v_p(f) = \frac{\partial f}{\partial x}(p)$$

or if $v = (0, 1)$ then we get

$$v_p(f) = \frac{\partial f}{\partial y}(p).$$

Theorem 2. *Suppose that the partial derivatives of* f *with respect to* x *and* y *exist everywhere and that they vary continuously. Let* $v = (v_1, v_2)$ *be a vector and* $p \in \mathbf{R}^2$ *be a point. Then*

$$v_p(f) = v_1 \frac{\partial f}{\partial x}(p) + v_2 \frac{\partial f}{\partial y}(p).$$

This theorem is a nice piece of analysis and therefore we won't prove it. You can tell it's nontrivial because it's not immediately obvious what continuity of the partial derivatives has to do with anything.

The directional derivatives are just linear combinations of the directional derivatives in the x- and y-directions. What does this mean geometrically, in terms of the graph of f?

In single-variable calculus, vectors in $\mathbf{R}$ have just one component v. The directional derivative of a function $g\colon \mathbf{R} \to \mathbf{R}$ along the vector v at x is $v_x(g) = vg'(x)$. As v varies this traces out a line with slope $g'(x)$, which is tangent to the graph of g when we translate it to $(x, g(x))$.

In two variables, the directional derivative of $f\colon \mathbf{R}^2 \to \mathbf{R}^2$ in the direction $v = (v_1, v_2)$ at $p = (x_0, y_0)$ is

$$v_p(f) = v_1 \frac{\partial f}{\partial x}(p) + v_2 \frac{\partial f}{\partial y}(p)$$

which traces out a plane as v_1 and v_2 vary. This plane is tangent to the graph of f when it is translated to the point $(x_0, y_0, f(x_0, y_0))$. In other words, the geometric content of the above theorem is that the graph of a function with continuous partial derivatives admits a *tangent plane*.

Example 3. *You cannot drop the assumption of continuity of partial derivatives. A counterexample would be the function*

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} \text{ when } (x,y) \neq (0,0) \\ 0 \text{ at } (x,y) = (0,0) \end{cases} .$$

Both partial derivatives exist everywhere. They vanish at the origin, but for example

$$\frac{\partial f}{\partial x} = \frac{y^3 - x^2 y}{(x^2+y^2)^2}$$

and along the y-axis this is $y^3/y^4 = 1/y$ which does not tend to zero as $y \to 0$, hence the partial derivatives are not continuous. The function is not even continuous when restricted to a line which is not $x = 0$ or $y = 0$ and hence the other directional derivatives don't even exist.

To get away from difficulties like this, we will only discuss functions with continuous first partial derivatives, also called $\mathcal{C}^1$-functions, or once continuously differentiable functions.

In this case we can define *total derivative* at p

$$d_p f\colon \mathbf{R}^2 \to \mathbf{R}$$

which encodes all directional derivatives. The total derivative eats a vector v and outputs the directional derivative in the v-direction at p, that is

$$d_p f(v) = v_p(f).$$

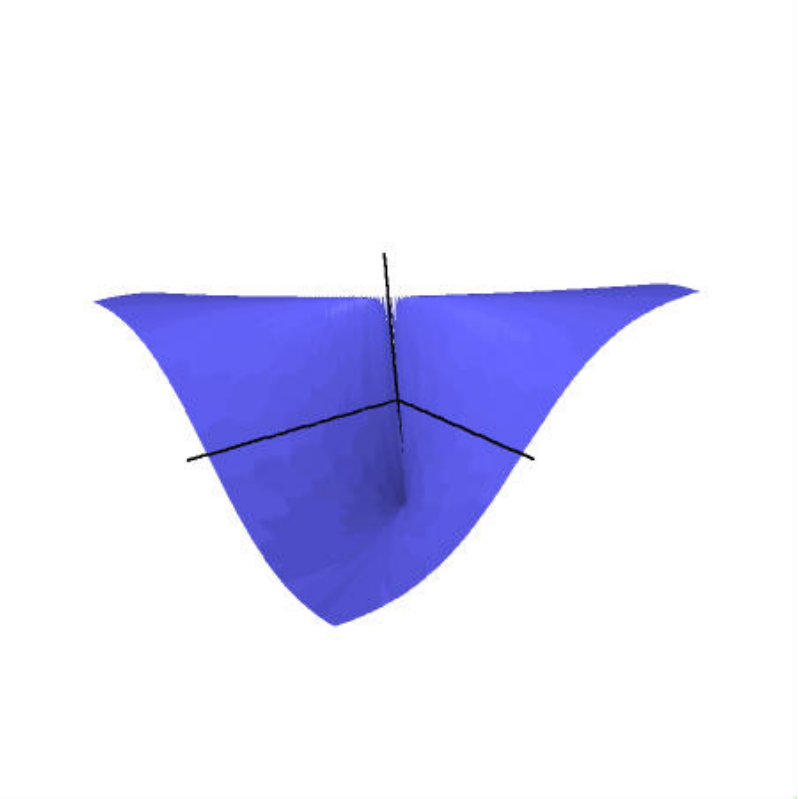


Figure 1.1: The graph of the function from Example 3 which has discontinuous partial derivatives and hence no tangent plane at the origin.

By Theorem 2, at each point this is a linear map

$$\begin{aligned} d_p f \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} &= (\partial_x f)v_1 + (\partial_y f)v_2 \\ &= \begin{pmatrix} \partial_x f & \partial_y f \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \end{aligned}$$

so we can write df as a row-vector (also known as a covector)

$$df = \begin{pmatrix} \partial_x f & \partial_y f \end{pmatrix}.$$

1.1.3 Linear approximation

The point of introducing the total derivative $d_p f \colon \mathbf{R}^2 \to \mathbf{R}$ is that it is a good linear (or "first order") approximation to the original function $f \colon \mathbf{R}^2 \to \mathbf{R}$ at the point p.

We first recall the situation in one dimension. Let $g\colon \mathbf{R} \to \mathbf{R}$ be a differentiable function of one variable. By the definition of the derivative we have the following fact

$$g(x_0 + \epsilon) = g(x_0) + \epsilon g'(x_0) + \epsilon\eta(\epsilon)$$

where $\eta(\epsilon) \to 0$ as $\epsilon \to 0$. In other words, $g(x_0)+\epsilon g'(x_0)$ is a good approximation of $g(x_0+\epsilon)$ for small ϵ. This is actually equivalent to the usual definition because by rearranging we have

$$\frac{g(x_0 + \epsilon) - g(x_0)}{\epsilon} = g'(x_0) + \eta(\epsilon)$$

Let $f\colon \mathbf{R}^2 \to \mathbf{R}$ be a function with continuous partial derivatives. Let (v_1, v_2) be a vector and $p = (x_0, y_0)$ a point. Since the above approximation holds in all directions and using Theorem 2 to express the directional derivative in the (v_1, v_2)-direction in terms of the partial derivatives, we get

$$f(x_0 + v_1, y_0 + v_2) = f(x_0, y_0) + v_1\partial_x f(x_0, y_0) + v_2\partial_y f(x_0, y_0) + |v|\eta(|v|)$$

where $\eta(\epsilon) \to 0$ as $\epsilon \to 0$. In other words,

$$f(x_0 + v_1, y_0 + v_2) = f(x_0, y_0) + d_p f(v) + |v|\eta(|v|).$$

Remark 4. *The total derivative $d_p f$ is a good approximation to f at p in the same way that the tangent plane to the graph* $\mathrm{Graph}(f)$ *of f at $(x_0, y_0, f(x_0, y_0))$ in $\mathbf{R}^3$ stays close to the graph itself in a neighbourhood of p. In fact, the tangent space at p to* $\mathrm{Graph}(f)$ *is the image $d_p f(\mathbf{R}^2)$, translated to the point $(x_0, y_0, f(x_0, y_0))$.*

1.1.4 The gradient

Another, equivalent way of packaging the information in the total derivative is to give the gradient ∇f of the function f. This is the *column vector*

$$(\nabla f)(p) = \begin{pmatrix} \frac{\partial f}{\partial x}(p) \\ \frac{\partial f}{\partial y}(p) \end{pmatrix}$$

(with more vertical entries if there are more coordinates). Note that by definition:

$$d_p f(v) = (\nabla f)(p) \cdot v.$$

Lemma 5. *The gradient $(\nabla f)(p)$ points in the direction of maximal increase of f at p. Moreover, its magnitude equals the directional derivative in the corresponding unit direction $(\nabla f)(p)/|(\nabla f)(p)|$.*

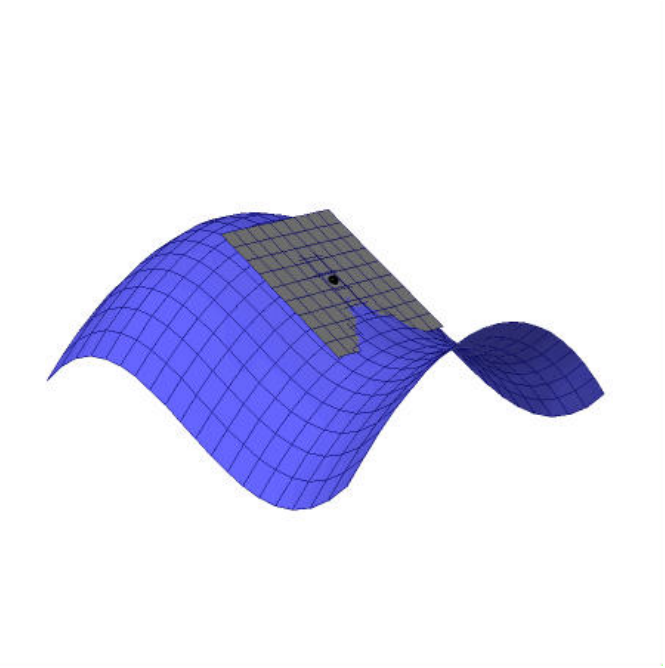


Figure 1.2: The graph of the total derivative $d_p f$ translated to the point $(p, f(p))$ is the tangent plane to Graph(f) at $(p, f(p))$.

Proof. Let's assume that p is not a critical point of f, otherwise the Lemma is obvious (because both $d_p f$ and $(\nabla f)(p)$ vanish). If v is a unit vector then $|d_p f(v)| = |(\nabla f)(p) \cdot v| = |(\nabla f)(p)||v| \cos(\theta)$ where θ is the angle between v and $(\nabla f)(p)$. This is clearly maximal when $\cos(\theta) = 1$, that is when $\theta = 0$. The directional derivative of f in the $(\nabla f)(p)/|(\nabla f)(p)|$-direction is

$$d_p f\left(\frac{(\nabla f)(p)}{|(\nabla f)(p)|}\right) = (\nabla f)(p) \cdot \frac{(\nabla f)(p)}{|\nabla f|} = |(\nabla f)(p)|.$$

□

1.1.5 The matrix of partial derivatives

In general for a vector-valued function $\mathbf{F} = (F_1, \ldots, F_m) \colon \mathbf{R}^n \to \mathbf{R}^m$ there are $n \times m$ derivatives which we write in an m-by-n matrix

$$d_p F = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(p) & \cdots & \frac{\partial F_1}{\partial x_n}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1}(p) & \cdots & \frac{\partial F_m}{\partial x_n}(p) \end{pmatrix}$$

In particular, for a change of variables $F\colon \mathbf{R}^2 \to \mathbf{R}^2$,

$$F(x,y) = (u(x,y), v(x,y))$$

we define the *Jacobian matrix*

$$\begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix}$$

The naturality of writing this as a matrix comes from the following theorem.

Theorem 6 (Chain rule)**.** *If $F\colon \mathbf{R}^n \to \mathbf{R}^m$ and $G\colon \mathbf{R}^m \to \mathbf{R}^\ell$ are differentiable functions then*

$$d_p(G \circ F) = d_{F(p)}G \circ d_pF$$

where $\circ$ denotes composition of functions on the left-hand side and matrix multiplication on the right.

We first make explicit note of a useful special case.

Corollary 7. *Suppose that $F\colon \mathbf{R}^2 \to \mathbf{R}^2$ is a change of coordinates, $(u,v) = F(x,y)$, and $G\colon \mathbf{R}^2 \to \mathbf{R}$ is a function. The function $G \circ F$ expresses $G(u,v)$ in terms of the coordinates (x,y) and we have*

$$\begin{pmatrix} \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial G}{\partial u} & \frac{\partial G}{\partial v} \end{pmatrix} \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix}$$

$$\text{i.e. } \frac{\partial G}{\partial x} = \frac{\partial G}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial G}{\partial v}\frac{\partial v}{\partial x},$$
$$\frac{\partial G}{\partial y} = \frac{\partial G}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial G}{\partial v}\frac{\partial v}{\partial y}.$$

Proof of Theorem 6. See `http://youtu.be/BtgBCPdT8rM`. Let v be a vector and ϵ a parameter which we assume takes on small values. We have

$$F(p+\epsilon v) = F(p) + \epsilon d_pF(v) + \epsilon|v|\eta(\epsilon|v|)$$

and

$$G(\mathbf{u}+\epsilon w) = G(\mathbf{u}) + \epsilon d_{\mathbf{u}}G(w) + \epsilon|w|\xi(\epsilon|w|)$$

for some functions $\eta, \xi\colon \mathbf{R} \to \mathbf{R}^m$ such that $|\eta(r)|$ and $|\xi(r)|$ tend to zero as $r \to 0$. Therefore

$$G(F(p+\epsilon v)) = G(F(x) + \epsilon d_pF(v) + \epsilon|v|\eta(\epsilon|v|))$$

so if we set $\mathbf{u} = F(p)$ and $w = d_pF(v) + |v|\eta(\epsilon|v|)$, we get

$$G(F(p+\epsilon v)) = G(F(p)) + \epsilon d_{F(p)}G \circ (d_pF(v) + |v|\eta(\epsilon|v|)) + \epsilon|w|\xi(\epsilon|w|)$$

which equals

$$G(F(p)) + \epsilon d_{F(p)}G \circ d_pF(v) + \epsilon \left(d_{F(p)}G(|v|\eta(\epsilon|v|)) + |w|\xi(\epsilon|w|) \right)$$

by linearity of $d_{F(p)}G$. The term

$$d_{F(p)}G(|v|\eta(\epsilon|v|)) + |w|\xi(\epsilon|w|)$$

goes to zero as $\epsilon \to 0$ and since this linear approximation determines our first derivative we know that

$$d_p(G \circ F) = d_{F(p)}G \circ d_pF$$

□

In case you find the level of abstraction in this proof intimidating, I suggest you try to prove the corollary explicitly in coordinates as it is stated. In case you find the proof fun, I suggest you go to Analysis 4 to see more where that came from.

1.1.6 Critical points

We are interested in maxima and minima of functions. I just want to recall:

Definition 8. *A function f has a* local maximum *(respectively* local minimum*) at p if there is a neighbourhood of p such that $f(p) \geq f(x)$ (respectively $f(p) \leq f(x)$) for all x in the neighbourhood.*

In one variable, we have the following theorem:

Theorem 9. *Suppose that $g\colon \mathbf{R} \to \mathbf{R}$ is a differentiable function with a local maximum or a local minimum at $x_0 \in \mathbf{R}$. Then $g'(x_0) = 0$.*

Geometrically, you drop a horizontal line down onto the graph of f until it hits. The point where it hits is clearly a local maximum and the horizontal line is tangent there.

Proof. Let's prove it for local maxima. Since g is differentiable it can be approximated by

$$g(x_0 + \epsilon) = g(x_0) + \epsilon g'(x_0) + \epsilon\eta(\epsilon)$$

where $\eta \to 0$ as $\epsilon \to 0$. If $g'(x_0) \neq 0$ then it has a sign, $\pm$. We consider small ϵ with the same sign. When the magnitude of ϵ is sufficiently small, $\eta(\epsilon)$ can be made arbitrarily small relative to $g'(x_0)$ and hence the error term $\epsilon\eta(\epsilon)$ can be made arbitrarily smaller in magnitude than $\epsilon g'(x_0)$.

In particular, if

$$g(x_0) + \epsilon g'(x_0) > g(x_0)$$

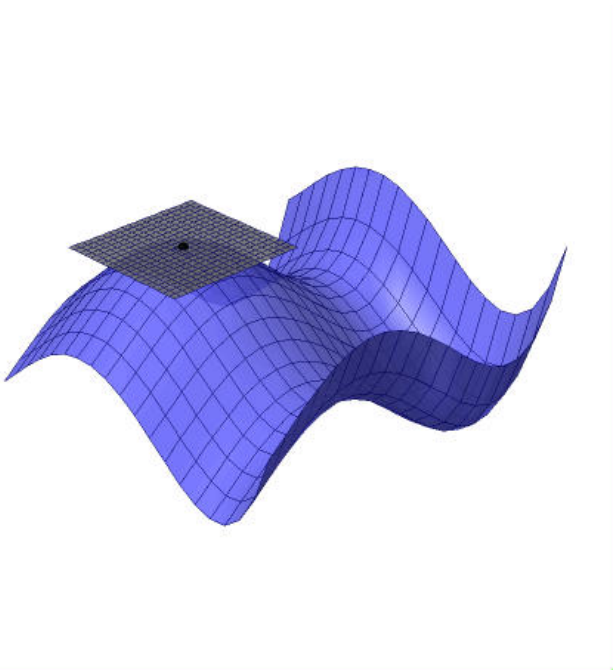


Figure 1.3: A function with a local maximum. Note that the tangent plane to the graph at the local maximum is horizontal.

for very small ϵ then

$$g(x_0 + \epsilon) = g(x_0) + \epsilon g'(x_0) + \epsilon\eta(\epsilon) > g(x_0)$$

because the error term is not big enough to change the inequality. Because we took ϵ to have the same sign as $g'(x_0)$ we certainly have $g(x_0) + \epsilon g'(x_0) > g(x_0)$ and hence this is a contradiction to local maximality of g at x_0. □

In two variables, the geometric idea is the same: you drop a horizontal plane down onto the graph and when it hits it is tangent (see Figure 1.3).

Theorem 10. *If $f\colon \mathbf{R}^2 \to \mathbf{R}$ is a $\mathcal{C}^1$-function with a local maximum or local minimum at $p = (x_0, y_0)$ then all directional derivatives vanish at p, in other words $d_p f = 0$.*

Proof. Suppose f has a local maximum or minimum at p. Let $\{(x_0 + tv_1, y_0 + tv_2) \ : \ t \in \mathbf{R}\}$ be the line through p in the direction $v = (v_1, v_2)$. Restrict f to this line; it still has a maximum or minimum at $t = 0$. Hence, by Theorem 9,

$$\left.\frac{d}{dt}\right|_{t=0} f(x_0 + tv_1, y_0 + tv_2) = 0$$

but this is precisely $v_p(f)$. □

Definition 11 (Critical point)**.** *Let $f\colon \mathbf{R}^2 \to \mathbf{R}$ be a $\mathcal{C}^1$-function. A point $p \in \mathbf{R}^2$ with $d_p f = 0$ is called a critical point. Equivalently, all directional derivatives of f vanish at p, or the tangent plane of the graph of f at p is horizontal.*

Not all critical points are maxima and minima. Even in one variable we have inflection points. The analogous critical points in two variables are the *saddle points*, see Figure 1.4.

1.2 Second derivatives

In one variable there is a sufficient condition for a critial point to be a local maximum (or a local minimum) in terms of its second derivatives. This is called the second derivative test and is proved using Taylor's theorem:

Theorem 12 (Taylor's theorem)**.** *Suppose that $g\colon \mathbf{R} \to \mathbf{R}$ is a twice differentiable function with derivatives g' and g''. Then*

$$g(x_0 + \epsilon) = g(x_0) + \epsilon g'(x_0) + \frac{\epsilon^2}{2} g''(x_0) + \epsilon^2 \eta(\epsilon)$$

for some function η which tends to zero as $\epsilon \to 0$.

Corollary 13 (Second derivative test)**.** *Suppose that $g\colon \mathbf{R} \to \mathbf{R}$ is a differentiable function and that g has a critical point at x_0. Suppose moreover that $g''(x_0) < 0$. Then g has a local maximum at x_0.*

Proof. From Taylor's Theorem:

$$\begin{aligned} g(x_0 + \epsilon) &= g(x_0) + \epsilon g'(x_0) + \frac{1}{2}\epsilon^2 g''(x_0) + \epsilon^2 \eta(\epsilon) \\ &= g(x_0) + \frac{1}{2}\epsilon^2 g''(x_0) + \epsilon^2 \eta(\epsilon) \end{aligned}$$

since $g'(x_0) = 0$ by the condition that x_0 is critical. Now by taking ϵ small enough, we can ensure that $|\eta(\epsilon)| < \frac{1}{2}|g''(x_0)|$ so if $g''(x_0) < 0$ then $g''(x_0) + \eta(\epsilon) < 0$ and hence

$$g(x_0 + \epsilon) < g(x_0)$$

for all small ϵ. Hence g has a local maximum at x_0. □

We will now develop the formalism of second partial derivatives in two variable calculus, prove the analogue of Taylor's theorem and explain the second derivative test for critical points in this context.

1.2.1 Definition

The second partial derivative

$$\frac{\partial^2 f}{\partial x^2}$$

of a function $f\colon \mathbf{R}^2 \to \mathbf{R}$ is defined by differentiating the function $\frac{\partial f}{\partial x}$ with respect to x, keeping y fixed. Similarly one can define partial derivatives

$$\frac{\partial^2 f}{\partial x \partial y}, \ \frac{\partial^2 f}{\partial y \partial x}, \ \frac{\partial^2 f}{\partial y^2}.$$

As before, to avoid analytical subtleties, we will assume that all partial derivatives exist and are continuous, a property which we will call $\mathcal{C}^2$. With this assumption we have the following

Theorem 14 (Symmetry of mixed partial derivatives)**.** *If $f\colon \mathbf{R}^2 \to \mathbf{R}$ is a $\mathcal{C}^2$-function then*

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

Given the definition

$$v_p(f) \equiv \left.\frac{d}{dt}\right|_{t=0} f(x_0 + tv_1, y_0 + tv_2)$$

of the directional derivative in the direction $v = (v_1, v_2)$ at $p = (x_0, y_0)$, the definition of a "second directional derivative" should be clear:

$$\left.\frac{d^2}{dt^2}\right|_{t=0} f(x_0 + tv_1, y_0 + tv_2).$$

Equivalently, consider the directional derivative $v_p(f)$ as a function of p,

$$v_\cdot(f)\colon \mathbf{R}^2 \to \mathbf{R}$$

and take its directional derivative in the v direction:

$$v_p(v_\cdot(f))$$

Henceforth we'll omit the subscript p, so this can be written as

$$v^2(f).$$

Of course you could also differentiate $v(f)$ in the w-direction for a different vector w. Taking v and w to be standard basis vectors, you end up with the partial derivatives

$$\frac{\partial^2 f}{\partial x^2}, \ \frac{\partial^2 f}{\partial x \partial y}, \ \frac{\partial^2 f}{\partial y^2}.$$

Lemma 15. *If $v = (v_1, v_2)$ then*

$$v^2(f) = v_1^2 \frac{\partial^2 f}{\partial x^2} + 2v_1 v_2 \frac{\partial^2 f}{\partial x \partial y} + v_2^2 \frac{\partial^2 f}{\partial y^2}.$$

Proof. We know that

$$v(f) = (v_1 \partial_x + v_2 \partial_y) f$$

so

$$\begin{aligned} v(v(f)) &= (v_1 \partial_x + v_2 \partial_y)(v_1 \partial_x + v_2 \partial_y) f \\ &= \left(v_1^2 \frac{\partial^2}{\partial x^2} + 2v_1 v_2 \frac{\partial^2}{\partial x \partial y} + v_2^2 \frac{\partial^2}{\partial y^2} \right) f \end{aligned}$$

by the binomial theorem. □

The generalisation to higher derivatives can be derived similarly and involves binomial coefficients.

1.2.2 The Hessian and Taylor's theorem

The expression

$$v_1^2 \frac{\partial^2 f}{\partial x^2}(p) + 2v_1 v_2 \frac{\partial^2 f}{\partial x \partial y}(p) + v_2^2 \frac{\partial^2 f}{\partial y^2}(p)$$

is a quadratic form in the coefficients of v. We can write it as

$$\begin{pmatrix} v_1 & v_2 \end{pmatrix} \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

The two-by-two matrix is called *the Hessian of f at p*, $\mathrm{Hess}_p(f)$. The quadratic form in the coefficients of v is $v^T \mathrm{Hess}_p(f) v$.

Theorem 16. *Suppose that $f \colon \mathbf{R}^2 \to \mathbf{R}$ is a $\mathcal{C}^2$-function. Then*

$$f(p+v) = f(p) + d_p f(v) + \frac{1}{2} v^T \mathrm{Hess}(f) v + |v| \eta(|v|)$$

for some function $\eta(r)$ which tends to zero as $r \to 0$.

1.2.3 The second derivative test

Theorem 17. *Suppose that $f \colon \mathbf{R}^2 \to \mathbf{R}$ is a $\mathcal{C}^2$-function and suppose moreover that p is a critical point of f, i.e. $d_p f = 0$. If both eigenvalues of the matrix $\mathrm{Hess}(f)$ are negative (respectively positive) then p is a local maximum (respectively local minimum).*

Proof. First notice that $\mathrm{Hess}_p(f)$ is a symmetric matrix. Therefore its eigenvalues are both real, say $\lambda_1, \lambda_2 \in \mathbf{R}$. Suppose they are both negative. We can diagonalise $\mathrm{Hess}_p(f)$ by an orthogonal matrix U:

$$U^T \mathrm{Hess}_p(f) U = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

Let $\mathbf{e}_1 = (1,0)$ and $\mathbf{e}_2 = (0,1)$ denote the basis vectors. Then $U\mathbf{e}_1$ is a λ_1-eigenvector for $\mathrm{Hess}_p(f)$. Taking the Taylor expansion we get

$$f(p + \epsilon U\mathbf{e}_1) = f(p) + \epsilon d_p f(U\mathbf{e}_1) + \frac{1}{2} e_1^T U^T \mathrm{Hess}_p(f) U\mathbf{e}_1 + \epsilon\eta(\epsilon)$$

Since p is a critical point, the first derivative term vanishes and we get

$$f(p + \epsilon U\mathbf{e}_1) = f(p) + \frac{1}{2} e_1^T U^T \mathrm{Hess}_p(f) U\mathbf{e}_1 + \epsilon\eta(\epsilon)$$

Using the fact that $U^T = U^{-1}$ (orthogonality of U), the second derivative term is

$$\begin{aligned} \frac{1}{2} e_1^T U^T \mathrm{Hess}_p(f) U\mathbf{e}_1 &= \frac{1}{2} \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \frac{1}{2}\lambda_1 \\ &< 0 \end{aligned}$$

and similarly for λ_2. The error term can be made very small when ϵ is very small, so that $f(p + \epsilon\mathbf{e}_1) < f(p)$. Similarly for $\mathbf{e}_2$ and indeed for any $v = v_1\mathbf{e}_1 + v_2\mathbf{e}_2$ we get

$$f(p + v_1 e_1 + v_2 e_2) = f(p) + \frac{1}{2}(\lambda_1 v_1^2 + \lambda_2 v_2^2) + \cdots$$

which is $< f(p)$ when $|v|$ is very small because λ_1, λ_2 are both negative. This proves that p is a local maximum and the same proof would demonstrate a local minimum in case both eigenvalues were positive. □

Corollary 18. *Suppose that $f\colon \mathbf{R}^2 \to \mathbf{R}$ is a $\mathcal{C}^2$-function and suppose moreover that p is a critical point of f. If*

$$\det(\mathrm{Hess}_p(f)) > 0$$

then p is either a local maximum or a local minimum of f. If moreover

$$\mathrm{Trace}(\mathrm{Hess}_p(f)) < 0, \ \textit{(respectively } > 0\textit{)}$$

then p is a local maximum (respectively local minimum).

Proof. Since $\mathrm{Hess}_p(f)$ can be diagonalised by an orthogonal matrix U,

$$\begin{aligned}\det(\mathrm{Hess}_p(f)) &= \det\left(U \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} U^T\right) \\ &= \det(U)\lambda_1\lambda_2 \det(U^T) \\ &= \lambda_1\lambda_2\end{aligned}$$

since $\det(U) = \det(U^T) = \pm 1$.

If $\det(\mathrm{Hess}_p(f)) > 0$ this means that λ_1 and λ_2 are nonzero and have the same sign. Hence Theorem 17 applies.

Of course, we cannot tell whether λ_1 and λ_2 are both positive or both negative without further information and hence we don't know whether p is a maximum or a minimum. In the case when λ_1 and λ_2 have the same sign, it suffices to compute the trace of $\mathrm{Hess}_p(f)$, since this is equal to $\lambda_1 + \lambda_2$ which has the same sign as both λ_1 and λ_2. This proves the theorem. □

Remark 19. *When there are eigenvalues equal to zero, the Taylor series argument doesn't work because the error term dominates the second order terms in certain directions. We call such critical points* degenerate. *They are more complicated.*

Definition 20. *A critical point p of a function $f\colon \mathbf{R}^2 \to \mathbf{R}$ is called* nondegenerate *if all its eigenvalues are nonzero. Equivalently, if* $\det(\mathrm{Hess}_p(f)) \neq 0$.

So what do things look like at a nondegenerate critical point if the determinant of the Hessian at a critical point is negative? That is, if one of the eigenvalues is positive and the other is negative? This is called a *saddle point*.

Example 21 (A saddle point)**.** *Consider the function*

$$f(x, y) = x^2 - y^2.$$

The origin is a critical point and the Hessian is

$$\mathrm{Hess}_0(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

with eigenvalues ± 1. We can see from Figure 1.4 that the function has a maximum along the y-direction and a minimum along the x-direction.

1.2.4 Higher dimensions

See `http://youtu.be/CNZOEPGKzcA`.

In higher dimensions the Hessian is a larger matrix, but symmetry of partial derivatives means that it is still symmetric. More complicated nondegenerate

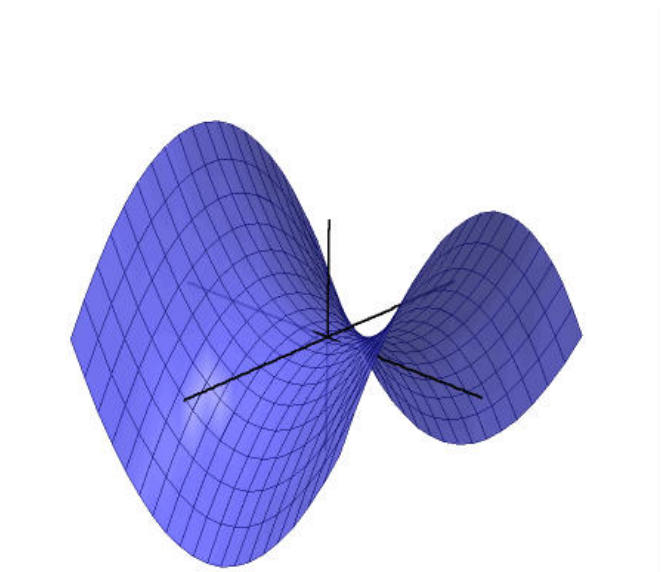


Figure 1.4: A function with a saddle point.

critical points are possible, but because we can always diagonalise a symmetric matrix, these critical points will always look like one of a finite list of models depending on how many positive and negative eigenvalues the Hessian has. These models are

$$\sum_{k=1}^{m} x_k^2 - \sum_{k=m+1}^{n} x_k^2$$

and are known as *quadratic* or *Morse singularities*.

Note that the diagonalisation argument still works: if p is a critical point of f and $u_1, \ldots, u_n$ is an orthonormal basis of eigenvectors for $\mathrm{Hess}_p(f)$ and $v = v_1 u_1 + \cdots + v_n u_+ n$ is a vector then the Taylor expansion is

$$f(p+v) = f(p) + \frac{1}{2}(\lambda_1 v_1^2 + \cdots + \lambda_n v_n^2) + \cdots,$$

so maxima are still points where all λ_k are negative and minima are still points where all λ_k are positive.

1.3 Constrained optimisation

1.3.1 The geometric idea...

We are often interested in solving a maximisation/minimisation problem with an extra constraint.

Example 22.
- *How close does the ellipse $x^2/a^2 + y^2/b^2 = 1$ get to the point p? This is quite easy: just parametrise the ellipse as $x = a\cos(t)$, $y = b\sin(t)$, write the distance between $(x(t), y(t))$ and p as a function of t and find and analyse the stationary points.*
- *How close does the surface $x^3 + y^3 + z^3 = 3xyz$ get to the point $(0, 0, 1)$? It's not so easy to find a convenient parametrisation of this cubic surface!*

For the sake of clarity we restrict attention to functions $f\colon \mathbf{R}^2 \to \mathbf{R}$ of two variables and, given a curve $C \subset \mathbf{R}^2$ we restrict the function f to C and obtain a new *constrained function* $h \equiv f|_C$. The question we want to answer is: what are the critical points of h?

There are two ways in which this can happen (see Figure 1.5):

- It is possible that a critical point of f happens to lie on C. For example, take $f(x, y) = x^2 - y^2$ and the curve $C = \{x = 0\}$. The restriction of f to C is the function $h(y) = -y^2$ (where y is the remaining coordinate on C). This curve passes through the saddle point at $(0, 0)$ and h has a maximum there.
- It is also possible that f has no critical points and yet $f|_C$ has some. For instance, take $f(x, y) = x$: since $\partial f/\partial x = 1$ this has no critical points. But if we restrict to the circle $C = \{x^2 + y^2 = 1\}$ then certainly $h(\theta) = \cos(\theta)$ has a maximum at $\theta = 0$ and a minimum at $\theta = \pi$ (where θ is the angular coordinate on the circle). Notice that these two critical points occur at the points where the level sets of f (the vertical lines) are tangent to the circle C. This is no coincidence...

Theorem 23 (See Figure 1.5)**.** *Let $f\colon \mathbf{R}^2 \to \mathbf{R}$ be a $\mathcal{C}^1$-function and $C \subset \mathbf{R}^2$ a curve in the plane. Let h denote the restriction $f|_C$. Then h has a critical point at $p \in C$ if and only if either*

- *f has a critical point at p or*
- *the level set $f^{-1}(f(p))$ and the level set C have the same tangent line at p.*

I should add the caveat that we need the C to be a nice smooth curve. In particular, we require that near every point $p \in C$ there is a $\mathcal{C}^1$-parametrisation $\gamma\colon (-T, T) \to C$ with $\gamma(0) = p$ and $\dot\gamma(0) \neq 0$. Some day you will see a theorem (maybe you already have) called the implicit function theorem which

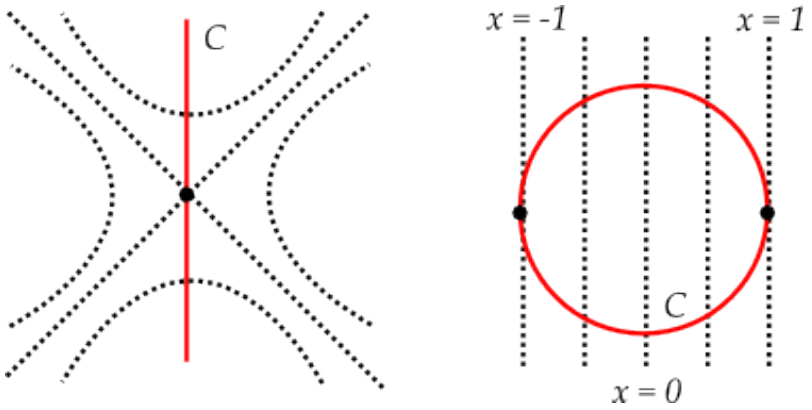


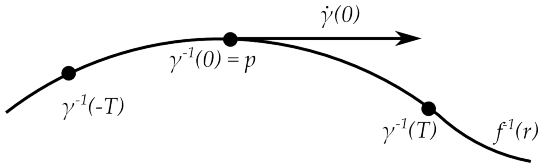
Figure 1.5: Examples illustrating Theorem 23. In one case (left), C runs through a critical point of f and inherits a critical point itself. In the other (right), the dashed contours of f are tangent to C at the two marked points where $h = f|_C$ has a minimum and a maximum

tells you that this parametrisation exists provided that C can be written as $g^{-1}(0)$ for a function g which has no critical points on the level set $g^{-1}(0)$. Crucially, we don't need to know the parametrisations to apply the theorem, we just need to know they exist. Similarly, the level sets of f admit local parametrisations near p provided p is not a critical point of f.

We start with a lemma to help us understand tangent lines to level sets:

Lemma 24. *If $f\colon \mathbf{R}^2 \to \mathbf{R}$ is a $\mathcal{C}^1$-function and $f^{-1}(r)$ is one of its level sets (assumed to contain no critical points) then the tangent line to $f^{-1}(r)$ at $p \in f^{-1}(r)$ is precisely the set of vectors v such that $d_p f(v) = 0$.*

Proof of Lemma 24. Take a parametrisation $\gamma\colon (-T, T) \to f^{-1}(r)$ with $\gamma(0) = p$ and $\dot{\gamma} \neq 0$. The tangent vectors we are looking for are precisely multiples of $\dot{\gamma}(0)$ so we want to show that these are annihilated by $d_p f$.



We know that

$$f \circ \gamma \equiv r$$

because the image of γ is (by definition) inside the level set $f^{-1}(r)$. Therefore $d_0(f \circ \gamma) = 0$. By the chain rule we get

$$0 = d_0(f \circ \gamma) = d_{\gamma(0)} f(\dot{\gamma}(0)).$$

So $d_p f$ annihilates $\dot\gamma(0)$, as required.

We also want to show that the kernel is *precisely* the set of vectors $\lambda\dot\gamma(0)$. Since p is not a critical point of f we know that there exists a vector v with $d_p f(v) \neq 0$. If u is any vector then $u = \lambda\dot\gamma(0) + \mu v$ for some $\lambda, \mu \in \mathbf{R}$ and hence

$$d_p f(u) = \mu.$$

This proves that if $d_p f(u) = 0$ then $\mu = 0$ and u is a multiple of $\dot\gamma(0)$. □

Proof of Theorem 23. Take $p \in C$ and let $\gamma\colon (-T, T) \to C$ be a parametrisation of a neighbourhood of p in the curve C. We are asking when p is a critical point of F, i.e. when $s = 0$ is a critical point of $s \mapsto f(\gamma(s))$. But $f(\gamma(s)) = (f \circ \gamma)(s)$ and by the chain rule

$$d_0(f \circ \gamma) = d_p f(\dot\gamma(0)).$$

This vanishes if and only if the direction $\dot\gamma(0)$ (which is tangent to C) is annihilated by $d_p f$. This certainly happens if p is a critical point of f, which is the first case of the theorem. If we assume that p is not a critical point of f then $d_p f(\dot\gamma(0)) = 0$ implies that $\dot\gamma(0)$ is tangent to the level set of f passing through p, by Lemma 24. □

1.3.2 ...and in practice

In practice we write $C = \{g = 0\}$ for some function g. By Lemma 24, the tangent line to a level set of f at p is the kernel of the linear map $d_p f$ (and similarly for g). If p is a critical point of $h = f|_{g^{-1}(0)}$ then, by Theorem 23, the level sets of f and g share a tangent line. That means that the two linear maps $d_p f$ and $d_p g$ have the same kernel and hence they are proportional:

$$d_p f = \lambda d_p g \text{ for some } \lambda.$$

We take λ as a new variable (the *Lagrange multiplier*) and consider the function

$$H(x, y, \lambda) = f(x, y) - \lambda g(x, y).$$

Varying with respect to x and y we get critical points when

$$d_p f - \lambda d_p g = 0,$$

i.e. when $d_p f$ and $d_p g$ are proportional. Varying with respect to λ, the critical point condition is

$$\frac{\partial H}{\partial \lambda} = -g(x, y) = 0.$$

In other words, critical points of H are triples (x, y, λ) such that

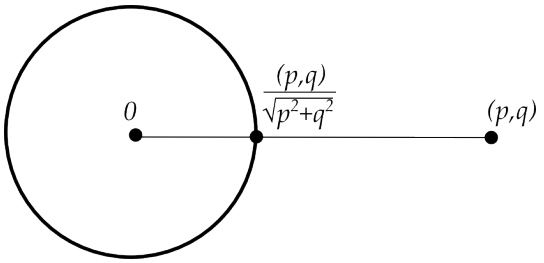


Figure 1.6: The setup for Example 25.

- $g(x, y) = 0$,
- (x, y) is a critical point of $F = f|_{g^{-1}(0)}$.

Example 25. *How close does the circle $x^2 + y^2 = 1$ get to the point (p, q)? We know the answer: just rescale (p, q) to have length one and you'll find the closest point to (p, q) on the circle. But for practice, let's work it out using Lagrange multipliers. Let $g(x, y) = x^2 + y^2 - 1$ and let $f(x, y) = (x-p)^2 + (y-q)^2$, which measures the (squared) distance to (p, q). In order to minimise f over $g^{-1}(0)$, we introduce a Lagrange multiplier λ and seek critical points of*

$$H(x, y, \lambda) = (x - p)^2 + (y - q)^2 + \lambda(x^2 + y^2 - 1)$$

Computing all the first derivatives we get

$$\begin{aligned} 2(x - p) + 2\lambda x &= 0 \\ 2(y - q) + 2\lambda y &= 0 \\ x^2 + y^2 &= 1 \end{aligned}$$

which gives

$$x = \frac{p}{1 + \lambda}, \qquad y = \frac{q}{1 + \lambda},$$

so

$$\frac{p^2}{(1 + \lambda)^2} + \frac{q^2}{(1 + \lambda)^2} = 1$$

or

$$\lambda = \sqrt{p^2 + q^2} - 1.$$

Substituting back gives

$$(x, y) = \frac{1}{\sqrt{p, q}}(p, q)$$

as we suspected.

Example 26. *Suppose you have enough money to buy four square metres of cardboard and you want to use it to make a lidless box. However, you're allowed to cut the cardboard into the desired shape before you buy it. How should you cut*

it to maximise the volume of the box? Well you can control the three dimensions of the box a, b, c and let's suppose that the lid would have area bc. The total surface area is therefore

$$2ab + 2ac + bc$$

and the volume is

$$f(a, b, c) = abc$$

Take $g(a, b, c) = 2ab + 2ac + bc - 4$ and introduce a Lagrange multiplier λ. We need to minimise

$$H(a, b, c, \lambda) = abc - \lambda(2ab + 2ac + bc - 4)$$

with respect to the four variables a, b, c, λ.

First, because the situation is symmetric with respect to b and c, we expect that the minimising configuration will have $b = c$. To prove this, we differentiate with respect to b and c and we see that, at a critical point,

$$0 = \frac{\partial H}{\partial b} = ac - \lambda(2a + c)$$
$$0 = \frac{\partial H}{\partial c} = ab - \lambda(2a + b)$$

This gives

$$(a - \lambda)c = 2\lambda a = (a - \lambda)b$$

so $b = c$ unless $a = \lambda$. If $a = \lambda$ then $ac - \lambda(2a + c) = 2a^2 = 0$ so $a = 0$, but we are assuming that our box has some *height. Therefore we can assume $b = c$.*

The problem is now to minimise

$$ab^2 - \lambda(4ab + b^2 - 4)$$

and the vanishing of partial derivatives with respect to a, b and λ gives

$$b^2 - 4\lambda b = 0$$
$$2ab - 4\lambda a - 2\lambda b = 0$$
$$4ab + b^2 = 4$$

Since $b \neq 0$, the first and third of these equations become

$$b = 4\lambda$$
$$a = \frac{4 - b^2}{4b}$$
$$= \frac{1 - 4\lambda^2}{4\lambda}$$

and hence the second becomes

$$1 = 12\lambda^2$$

Hence the solution is

$$a = 1/\sqrt{3}, \qquad b = 4/\sqrt{3}.$$

Chapter 2

Calculus of variations

Up to now we have been interested in a very limited class of optimisation problems: allowing ourselves a finite set of variables (usually one, two or three) and maximising a function over these. Things become much more interesting when we have an infinite amount of freedom to vary. For example, we might be interested in all paths betwen two points, or all surfaces in space with a given boundary curve, and we might be interested in minimising length, respectively area. We will first deal with a classic example: proving that a line between two points in the plane is the shortest path joining them. Then we will move on to the more general theory, illustrating it with a plethora of examples.

A very nice introduction to these ideas is provided by the Feynman lecture on the principle of least action.

2.1 Straight lines are shortest paths

I want to convince you that a straight line is the shortest path between two points in the plane. Maybe you don't need much convincing of this fact, but did you ever really think about what this means? I'm saying that if you fix two points and consider the space of all possible paths between them (which is of course an incomprehensibly large, in fact infinite dimensional, space) you can define a functional on this infinite-dimensional space of paths which takes each path to its length and this functional has a global minimum at the path given by a straight line. It suddenly seems like a more intimidating task to really convince ourselves that this is true, and it seems more surprising that our human intuition picked up on this without our noticing.

2.1.1 Paths and length

See `http://youtu.be/bM_klC-oAzg`. It seems like the key step in proving the statement is in understanding what we mean by a path. Let us fix two points A and B in the plane with coordinates (A_1, A_2) and (B_1, B_2). By a path between them we mean a map

$$\gamma\colon [0,1] \to \mathbf{R}^2$$

such that $\gamma(0) = A$ and $\gamma(1) = B$. Alternatively, we can project $\gamma(t)$ to the two coordinate axes and think of it as a pair of functions $\gamma(t) = (\gamma_1(t), \gamma_2(t))$ satisfying $\gamma_1(0) = A_1$, $\gamma_1(1) = B_1$, etc. Now we don't want our paths to jump around discontinuously in the plane so we'll certainly require the two functions γ_1 and γ_2 to be continuous. In fact, even to define the length of a path we're going to need to require slightly more:

- we need γ_1 and γ_2 to be differentiable and for the derivatives to be continuous;
- we also want the derivatives $\dot{\gamma}_1$ and $\dot{\gamma}_2$ never vanish simultaneously, in other words the vector $\dot{\gamma}$ is never allowed to vanish.

I'll explain why we need these when we need them.

When we try and define the length of a curved path, the natural thing to do is to zoom in very closely to the curve and recall from Taylor's theorem that on very small scales the path is well-approximated by a line. We can imagine taking finer and finer polygonal approximations of our path and defining the length to be the limit of the lengths of the polygons. This is of course the same as defining length by an integral along the curve and the infinitesimal length along the curve (the infinitesimal arc-length) is just dt times the length of the vector $\dot{\gamma}$ where the dot denotes time-differentiation. Using Pythagoras, this is

$$L(\gamma) \equiv \int_0^1 \sqrt{\dot{\gamma}_1^2 + \dot{\gamma}_2^2}\, dt$$

If we wanted to be wholly rigorous we would first show that the definition by polygonal approximations gave a well-defined number and then prove that this number could be computed as the above integral. Instead, I'll take the shortcut of defining the length of a differentiable path to be the above integral.

Remark 27. *Note that here we need the path to be differentiable to even write down the integrand. We also need the integral to be well-defined; if the derivative of γ is continuous then so is the integrand and we know how to integrate continuous functions on a closed interval (Riemann integration from Analysis I/II).*

2.1.2 Straight lines minimise action

Actually, the square root really causes us headaches, so I want to get rid of it for the moment and define the *action* of a path to be the integral

$$\alpha(\gamma) \equiv \int_0^1 \left(\dot{\gamma}_1^2 + \dot{\gamma}_2^2\right) dt.$$

We'll prove

Proposition 28. *If γ is a straight line between A and B and δ is another path between A and B then*

$$\alpha(\gamma) \leq \alpha(\delta)$$

with equality if and only if $\delta \equiv \gamma$.

In other words, straight lines minimise action. We'll see later how to deduce the claim about length from this claim about action.

Proof. Define $\epsilon(t) = \delta(t) - \gamma(t)$.

$$\begin{aligned}
\alpha(\delta) &= \int_0^1 \left(|\dot{\gamma} + \dot{\epsilon}|^2\right) dt \\
&= \int_0^1 \left(|\dot{\gamma}|^2 + 2\dot{\gamma} \cdot \dot{\epsilon} + |\dot{\epsilon}|^2\right) dt \\
&= \alpha(\gamma) + \alpha(\epsilon) + 2\int_0^1 \dot{\gamma} \cdot \dot{\epsilon} dt
\end{aligned}$$

Now because γ is linear, $\dot{\gamma}$ is constant and hence the final term is

$$2\dot{\gamma} \cdot \int_0^1 \dot{\epsilon} dt = 2\dot{\gamma} \cdot [\epsilon(t)]_{t=0}^1$$

by the fundamental theorem of calculus. But because $\epsilon(0) = \epsilon(1) = 0$ (remember the endpoints of the path are fixed) this term vanishes. Now we see that

$$\alpha(\delta) = \alpha(\gamma) + \alpha(\epsilon) \geq \alpha(\gamma)$$

with equality if and only if $\epsilon \equiv 0$, if and only if $\delta \equiv \gamma$. □

The key idea here was to integrate by parts (i.e. apply the fundamental theorem of calculus and use the fact that γ satisfies some highly restrictive equation ($\dot{\gamma}$ constant) to get rid of the term linear in ϵ. We were also lucky enough that the remaining term was strictly positive, which is what allows us to prove this incredibly strong result about global minimation of action by lines in Euclidean space. It's not usually the case in such variational arguments that we can control the higher-order terms in ϵ.

2.1.3 Straight lines minimise length

Now we want to argue that straight lines are in fact shortest. If I give you a differentiable path γ you can reparametrise it. To do this, take your favourite differentiable, monotonically strictly increasing function $\phi\colon [0,1] \to [0,1]$ and form the composition $\delta = \gamma \circ \phi$, in other words

$$\delta(t) = \gamma(\phi(t))$$

This gives a new path (differentiable by the chain rule!). Reparametrising doesn't change the length of a path but it can certainly change the action.

Proposition 29. *Any path can be reparametrised so that its action is equal to the square of its length.*

Proof. The parametrisation is defined as follows.

- Let
$$s(t) = \int_0^t |\dot{\gamma}| dt$$
be the arc-length after time t. Provided that $\dot{\gamma}$ is never zero, s is a monotonically increasing function of t.
- We normalise and consider the function $\phi(t) = s(t)/L(\gamma)$ which measures what proportion of the arc-length has ben traversed after time t. It is easy to see that $\phi(t)$ is differentiable since, by the fundamental theorem of calculus, its derivative is
$$\frac{d\phi}{dt}(t) = |\dot{\gamma}(t)|/L(\gamma).$$
Remember we are assuming that $|\dot{\gamma}| \neq 0$ everywhere.
- Since ϕ is monotonically increasing, it is a bijection and hence it has an inverse ϕ^{-1} which takes a input a number $\lambda \in [0,1]$ and outputs the time t at which γ has traversed an arc-length $\lambda L(\gamma)$. We will need the following fact about ϕ^{-1}:

Lemma 30. *If $\phi\colon [0,1] \to [0,1]$ is a once continuously-differentiable function whose derivative is always positive then it admits a continuously-differentiable inverse ϕ^{-1} whose derivative is equal to*

$$\frac{d\phi^{-1}}{dt} = 1 \Big/ \frac{d\phi}{dt}\,.$$

- Therefore ϕ^{-1} gives us a reparametrisation. What does the curve $\delta = \gamma \circ \phi^{-1}$ do? At time $t = 1/2$ it moves to the point on γ which is exactly halfway along (as measured by arc-length).

Now the reparametrised path $\delta = \gamma \circ \phi^{-1}$ has

$$\begin{aligned}\left|\frac{d\delta}{dt}\right| &= \left|\frac{d\gamma}{d\phi}\right| \frac{d\phi^{-1}}{dt} \\ &= L(\gamma)\frac{d\phi}{dt}\Big/\frac{d\phi}{dt} \\ &= L(\gamma)\end{aligned}$$

and hence its action is

$$\begin{aligned}\int_0^1 |\dot{\delta}(t)|^2 dt &= \int_0^1 L(\gamma)^2 dt \\ &= L(\gamma)^2.\end{aligned}$$

□

Now it is clear that the action of a straight line is $L(\gamma)^2$. This plus the previous proposition tells us that the length of a straight line is strictly less than the length of any other path joining the points!

2.2 The Euler-Lagrange equations

2.2.1 Key steps reviewed

Let us recall the key steps from the last section:

- Define some (infinite-dimensional) function space X; in that section it was the space of paths connecting A and B.
- Define a functional $F\colon X \to \mathbf{R}$ on that space; in that section it was the action functional.
- Take a supposed critical point, $\gamma \in X$ of the functional; in that section it was the straight line segment.
- Compute $F(\gamma + \epsilon)$, where $\gamma + \epsilon$ is a small variation of γ. Usually you can only compute this to first order in ϵ, in other words you can only usually compute the quantity we think of as the directional derivative of F in the ϵ-direction. We call this the *variation in F associated to the variation ϵ of γ* (this is where the name *variational calculus* comes from).
- Wherever the derivative of ϵ occurs in the resulting integral, integrate by parts.

Let's do this again for the action of a path:

$$\alpha(\gamma) \equiv \int_0^1 \left(\dot{\gamma}_1^2 + \dot{\gamma}_2^2\right) dt$$

We get

$$\begin{aligned}\alpha(\gamma+\epsilon) &= \int_0^1 \left((\dot\gamma_1+\dot\epsilon_1)^2+(\dot\gamma_2+\dot\epsilon_2)^2\right)dt \\ &= \alpha(\gamma) + 2\int_0^1 (\dot\gamma_1\dot\epsilon_1+\dot\gamma_2\dot\epsilon_2)\,dt + \mathcal{O}(\epsilon^2)\end{aligned}$$

so the directional derivative of α in the ϵ-direction at γ is

$$\begin{aligned}d_\gamma\alpha(\epsilon) &= \int_0^1 (\dot\gamma_1\dot\epsilon_1+\dot\gamma_2\dot\epsilon_2)\,dt \\ &= [\dot\gamma_1\epsilon_1+\dot\gamma_2\epsilon_2]_0^1 - \int_0^1 (\ddot\gamma_1\epsilon_1+\ddot\gamma_2\epsilon_2)\,dt\end{aligned}$$

where we have integrated by parts to get rid of $\dot\epsilon$s. Note that the boundary term vanishes because $\epsilon(0)=\epsilon(1)=0$.

Recall that a critical point is somewhere that all directional derivatives vanish. So the condition that γ is a critical point of α is just

$$d_\gamma\alpha(\epsilon) = 0 \text{ for all } \epsilon$$

i.e.

$$-\int_0^1 \ddot\gamma\cdot\epsilon dt = 0 \text{ for all } \epsilon.$$

We now state the

Theorem 31 (Fundamental theorem of the calculus of variations)**.** *Suppose that $y\colon [0,1]\to\mathbf{R}^n$ is a vector-valued function. If $\int_0^1 y(t)\cdot\epsilon(t)dt = 0$ for all $\mathcal{C}^1$-functions ("variations") $\epsilon\colon[0,1]\to\mathbf{R}^n$ then $y\equiv 0$.*

In particular this implies that $\ddot\gamma\equiv 0$, in other words the components of γ have to be linear function of t. In other words, γ is a linearly parametrised straight line. So even if we hadn't known the answer was a line to begin with, we could have worked it out (provided we could solve the differential equation $\ddot\gamma = 0$).

Remark 32. *Notice that the equation we obtained above is a second-order equation ($\ddot\gamma = 0$). This will hold in general because of the step where we integrate by parts. This means that we should always assume our function γ is twice continuously differentiable ($\mathcal{C}^2$) instead of just $\mathcal{C}^1$.*

2.2.2 The fundamental theorem of the calculus of variations

In this section we will prove Theorem 31.

Proof. Suppose, to the contrary, that there is a $t_0 \in [0,1]$ with $y(t_0) \neq 0$. We may as well assume that the component $y_1(t_0) > 0$. Because $y_1(t_0) > 0$, we know that $y_1(t) > 0$ for all t in some small interval $(t_0 - \delta, t_0 + \delta)$.

Define a "bump function" $f\colon [0,1] \to \mathbf{R}$ which is

- $\mathcal{C}^1$,
- nonnegative everywhere and positive at t_0,
- and vanishes outside the interval $(t_0 - \delta, t_0 + \delta)$.

Such functions are quite easy to construct. For instance

$$f(t) = \begin{cases} \exp\left(\frac{1}{(t-t_0)^2 - \delta^2}\right) & \text{if } t \in (t_0 - \delta, t_0 + \delta) \\ 0 & \text{otherwise} \end{cases}$$

will do!

Now consider the function $\epsilon\colon [0,1] \to \mathbf{R}^n$ given by

$$\epsilon(t) = (f(t), 0, \ldots, 0)$$

Integrating this against y gives

$$\int_0^1 y(t) \cdot \epsilon(t) dt = \int_{t_0 - \delta}^{t_0 + \delta} f(t) y_1(t) dt > 0$$

which contradicts the assumption that $\int_0^1 y(t) \cdot \epsilon(t) dt = 0$ for all ϵ. □

2.2.3 The Euler-Lagrange equation

Suppose we are interested in a situation where the space X is a space of $\mathcal{C}^2$-functions $\phi\colon [a,b] \to \mathbf{R}$ of one variable, t. Suppose moreover that the values of ϕ are specified at the endpoints of the interval $[a,b]$,

$$\phi(a) = A, \ \phi(b) = B.$$

The other functions in the space X can be written

$$\phi + \epsilon$$

for some variation $\epsilon\colon [a,b] \to \mathbf{R}$ with $\epsilon(a) = \epsilon(b) = 0$.

Let L be a function of three variables

$$L(p,q,r)$$

and suppose that the functional we want to minimise is

$$F(\phi) = \int_a^b L\left(t, \phi(t), \dot{\phi}(t)\right) dt,$$

in other words the integrand depends only on t, ϕ and its first derivative. For instance, the integrand in the length and area functionals above only depended on $\dot{\phi}$. We call the function L *the Lagrangian* of our problem.

Now perturb ϕ by a variation ϵ and expand the integrand to first order in ϵ using the chain rule:

$$L(t, \phi(t) + \epsilon(t), \dot{\phi}(t) + \dot{\epsilon}(t)) = L\left(t, \phi(t), \dot{\phi}(t)\right) + \epsilon(t)\frac{\partial L}{\partial p}\left(t, \phi(t), \dot{\phi}(t)\right) + \\ + \dot{\epsilon}(t)\frac{\partial L}{\partial q}\left(t, \phi(t), \dot{\phi}(t)\right) + \mathcal{O}(\epsilon^2)$$

Integrating we get

$$\begin{aligned} F(\phi + \epsilon) &= \int_a^b L\left(t, \phi + \epsilon, \dot{\phi} + \dot{\epsilon}\right) dt \\ &= F(\phi) + \int_a^b \frac{\partial L}{\partial q}\left(t, \phi, \dot{\phi}\right)\epsilon(t)dt + \int_a^b \frac{\partial L}{\partial r}\left(t, \phi, \dot{\phi}\right)\dot{\epsilon}(t) + \mathcal{O}(\epsilon^2) \\ &= F(\phi) + \int_a^b \left(\frac{\partial L}{\partial q}\left(t, \phi(t), \dot{\phi}(t)\right) - \frac{d}{dt}\left(\frac{\partial L}{\partial r}\left(t, \phi(t), \dot{\phi}(t)\right)\right)\right)\epsilon(t)dt + \\ &\quad + \left[\frac{\partial L}{\partial r}\epsilon\right]_a^b + \mathcal{O}(\epsilon^2) \end{aligned}$$

where we have integrated by parts and picked up a boundary term. The boundary term vanishes because $\epsilon(a) = \epsilon(b) = 0$. From this we see that the directional derivative of F in the ϵ-direction at ϕ is

$$\int_a^b \left(\frac{\partial L}{\partial q}\left(t, \phi(t), \dot{\phi}(t)\right) - \frac{d}{dt}\left(\frac{\partial L}{\partial r}\left(t, \phi(t), \dot{\phi}(t)\right)\right)\right)\epsilon(t)dt.$$

If ϕ is an critical point then all directional derivatives vanish, which means that the above integral vanishes *for all* ϵ. By Theorem 31 this means that

$$\frac{\partial L}{\partial q}\left(t, \phi(t), \dot{\phi}(t)\right) - \frac{d}{dt}\left(\frac{\partial L}{\partial r}\left(t, \phi(t), \dot{\phi}(t)\right)\right) = 0$$

which is now a second-order differential equation for ϕ. Note that we have written $\frac{d}{dt}$ and not $\frac{\partial}{\partial t}$ because the object being differentiated is actually just a function of one variable, t.

Because $L(p, q, r)$ is often written $L(t, \phi, \dot{\phi})$ the q- and r-derivatives are often written

$$\frac{\partial L}{\partial \phi} \text{ and } \frac{\partial L}{\partial \dot{\phi}}$$

which looks extremely confusing first time you see it, because how can ϕ and $\dot{\phi}$ be independent variables? You should just think of it as convenient shorthand

for the equation I've just written, and henceforth we'll adopt this notation too. The Euler-Lagrange equation is then

$$\boxed{\frac{\partial L}{\partial \phi} = \frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}}} \tag{2.1}$$

and that's what you'll always see written.

2.2.4 Beltrami identity

A second-order differential equation is usually not as easy as a first-order differential equation. If the Lagrangian $L(p, q, r)$ is independent of p then it turns out there is a first-order equation we can use instead of the Euler-Lagrange equation.

Theorem 33. *Suppose $L(p, q, r)$ has no p-dependence (i.e. $\partial L/\partial p = 0$). If ϕ satisfies the associated Euler-Lagrange equation then*

$$L\left(t, \phi(t), \dot{\phi}(t)\right) - \dot{\phi}(t)\frac{\partial L}{\partial r}\left(t, \phi(t), \dot{\phi}(t)\right) = C$$

for some constant C. This is called the Beltrami identity.

Of course, this is usually written

$$\boxed{L - \dot{\phi}\frac{\partial L}{\partial \dot{\phi}} = C.}$$

Proof. Throughout the proof, we write

$$L,\ \partial L/\partial p,\ \text{etc.}$$

instead of

$$L\left(t, \phi(t), \dot{\phi}(t)\right),\ \frac{\partial L}{\partial p}\left(t, \phi(t), \dot{\phi}(t)\right),\ \text{etc.}$$

Using the chain rule, we get

$$\frac{d}{dt}\left(L - \dot{\phi}(t)\frac{\partial L}{\partial r}\right) = \frac{\partial L}{\partial p} + \dot{\phi}(t)\frac{\partial L}{\partial q} + \ddot{\phi}(t)\frac{\partial L}{\partial r} - \ddot{\phi}(t)\frac{\partial L}{\partial r} - \dot{\phi}(t)\frac{d}{dt}\left(\frac{\partial L}{\partial r}\right)$$

By assumption $\partial L/\partial p = 0$ so the first term on the right-hand side vanishes. The two terms with $\ddot{\phi}$ cancel and we are left with

$$\frac{d}{dt}\left(L - \dot{\phi}(t)\frac{\partial L}{\partial r}\right) = \dot{\phi}(t)\frac{\partial L}{\partial q} - \dot{\phi}(t)\frac{d}{dt}\left(\frac{\partial L}{\partial r}\right).$$

The right-hand side is now clearly recognisable as $\dot{\phi}(t)$ times the Euler-Lagrange operator. Since we're assuming that ϕ satisfies the Euler-Lagrange equation, we see that $\frac{d}{dt}\left(L - \dot{\phi}(t)\frac{\partial L}{\partial r}\right) = 0$, so

$$L\left(t, \phi(t), \dot{\phi}(t)\right) - \dot{\phi}(t)\frac{\partial L}{\partial r}\left(t, \phi(t), \dot{\phi}(t)\right) = C$$

for some constant C, as required. □

When you study classical mechanics in the Lagrangian/Hamiltonian setting (Analytical Dynamics, next term) you will be able to understand this as a statement of conservation of energy for an autonomous Hamiltonian system.

Remark 34. *It seems strange that we can replace the second-order Euler-Lagrange equation with its two boundary conditions $\phi(a) = A$, $\phi(b) = B$ by a first-order equation with two boundary conditions (usually we only have the freedom to fix one boundary condition for a first-order equation). The point is that the Beltrami has an undetermined constant, C, which we can fix using the second boundary condition.*

2.2.5 Vector-valued functions

We often come across problems where we have to optimise over a space of vector-valued functions. For instance, showing that a straight line is shortest involved looking at two-component functions $(\gamma_1(t), \gamma_2(t))$. In this case there is an Euler-Lagrange equation for each component. Suppose that the vector-valued function is $(f_1(t), \ldots, f_n(t))$ and that

$$L(t, f_1(t), \ldots, f_n(t), \dot{f}_1(t), \ldots, \dot{f}_n(t))$$

is the Lagrangian. Then the Euler-Lagrange equations are

$$\begin{aligned} \frac{\partial L}{\partial f_1} &= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{f}_1}\right) \\ &\vdots \\ \frac{\partial L}{\partial f_n} &= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{f}_n}\right) \end{aligned}$$

This is easy to see by taking variations individually. For example, the variation $(f_1 + \epsilon_1, f_2, \ldots, f_n)$ gives rise to the first equation. In fact you could think of this as like the "partial derivative" of the functional in the f_1-direction (very loosely speaking).

Problem 35. *If $\partial L/\partial t = 0$ then a Beltrami identity holds:*

$$L - \sum_{k=1}^{n} \dot{f}_k \frac{\partial L}{\partial \dot{f}_k} = C.$$

2.2.6 A caveat

When we do these basic variational calculations all we are doing is finding the critical points of a particular functional. It is possible to do a second-derivative test to figure out if these are local maxima or minima, however most of the time we want to know when something is a *global maximum or minimum*, i.e. it's the biggest/smallest *of all possibilities*. The methods we are developing are purely local and won't tell us such global information, in the same way that ordinary calculus won't tell you global information for an ordinary function. The theorems we prove therefore have the form: "Assume a global minimum x of the functional $F\colon X \to \mathbf{R}$ were to exist. Then x would be a straight line/circle/catenary..."

Problem 36. *To illustrate the difficulty (in the finite-dimensional setting), sketch for me the graph of a function $f\colon \mathbf{R}^2 \to \mathbf{R}$ which admits a local minimum at the origin but has no global minimum.*

There *is* a set of harder analytical techniques which belong to the calculus of variations which allow one to prove the existence of global maximisers/minimisers. The idea is roughly the following. Suppose the functional $F\colon X \to \mathbf{R}$ you are interested in is bounded (say from below). Take a sequence $x_k \in X$ such that $F(x_k)$ tends to the infimum $\inf_{x\in X} F(x)$. Try to find a convergent subsequence. This subsequence may not have a limit in the space X - if the space X consists of $\mathcal{C}^2$-functions then maybe there is some loss of differentiability in the limit. Nonetheless, the subsequence has a limit in some slightly larger space $\overline{X}$ (using something like the Arzela-Ascoli theorem) and is a global minimum there. This implies that it satisfies a weak form of the Euler-Lagrange equation (weak in the sense that it may not be differentiable and hence the Euler-Lagrange equation doesn't even make sense for it). Now apply some kind of regularity theory to prove that the limit is actually as smooth as you want it to be, which implies that the global minimum exists somewhere in X. You will hopefully meet such compactness and regularity theorems in future functional analysis/calculus of variations/geometric analysis courses.

Of course in the examples we care about in this course, this technology proves the existence results we want. In the words of Hilbert:

> Every problem in the Calculus of Variations has a solution, provided the word 'solution' is suitably understood.

2.3 Examples

2.3.1 The catenoid

Let $\phi\colon [a,b] \to \mathbf{R}$ be a function with $\phi(a) = A$, $\phi(b) = B$ and $\phi(x) > 0$ for all $x \in [-1,1]$. Consider the graph of this function inside $\mathbf{R}^2$

$$\{(x,z) \in \mathbf{R}^2 \ : \ z = \phi(x)\}$$

and consider this $\mathbf{R}^2$ as the (x,z)-plane in $\mathbf{R}^3$. Rotate the graph around the x-axis and consider the surface it traces out. This is called a *surface of revolution*.

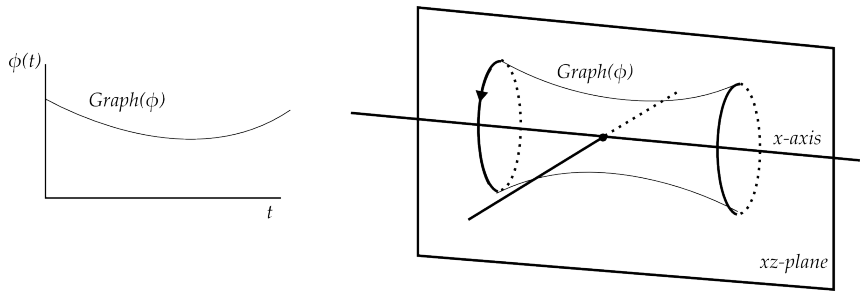


Figure 2.1: A surface of revolution, obtained by revolving a graph around an axis.

Lemma 37. *The area of this surface of revolution can be written as an integral*

$$2\pi \int_a^b \phi(t)\sqrt{1+\dot{\phi}(t)^2}dt.$$

Proof. See `http://youtu.be/p_xiCrZz_IU`. In order to define the area of the surface of revolution, we approximate it as follows. Take a decomposition of the interval $[a,b]$ into n pieces of width $\epsilon = (b-a)/n$. Let $t_k = k\epsilon$, $k = 0,\ldots,n$. We replace the surface by a union of *conical frustra* F_k, $k = 0,\ldots,n-1$, namely over the interval $[t_k, t_{k+1}]$ we define F_k to be the surface of revolution obtained by revolving a segment of tangent line to the graph of ϕ at $(t_k, \phi(t_k))$ around the x-axis. This frustum looks like a piece of a conical surface connecting a circle C_{t_k} and a circle D_{t_k} of radius Q_k and R_k respectively.

The surface area of this frustum is given by[1]

$$\text{area}(F_k) = \pi d_k(Q_k + R_k)$$

where d_k is the distance along the surface of the frustum between the two end circles. It is clear from the picture that $Q_k = \phi(t_k)$ and, since the radius

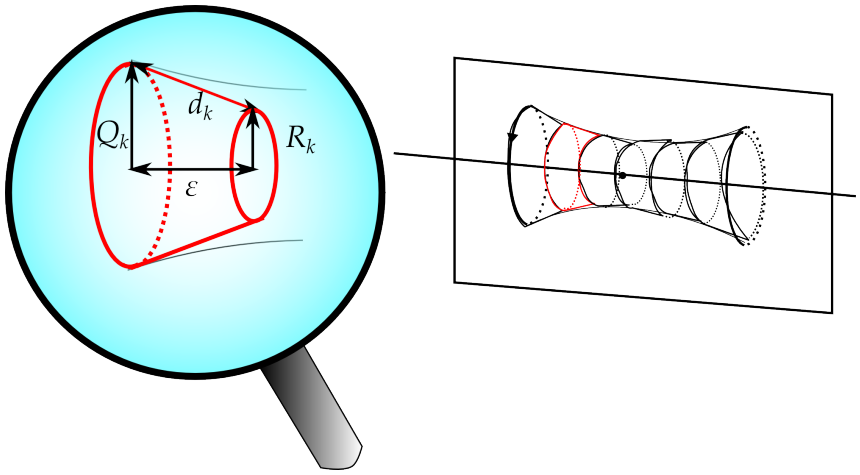


Figure 2.2: Approximating a surface of revolution by smaller and smaller cylindrical frusta.

R_k is obtained by moving from $(t_k, \phi(t_k))$ along a line of slope $\dot{\phi}(t_k)$ for time ϵ, we know that $R_k = Q_k + \dot{\phi}(t_k)\epsilon$. We can compute d_k from Pythagoras's theorem

$$d_k = \sqrt{\epsilon^2 + \epsilon^2 \dot{\phi}(t_k)^2} = \epsilon\sqrt{1 + \dot{\phi}(t_k)^2}$$

so the area of the frustum F_k is

$$\pi\epsilon\sqrt{1 + \dot{\phi}(t_k)^2}\left(2\phi(t_k) + \epsilon\dot{\phi}(t_k)\right).$$

The area of the whole surface is defined as the limit

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} \text{area}(F_k)$$

which is

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} \pi\epsilon\sqrt{1 + \dot{\phi}(t_k)^2}\left(2\phi(t_k) + \epsilon\dot{\phi}(t_k)\right)$$

Let us expand this:

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} 2\pi\epsilon\sqrt{1 + \dot{\phi}(t_k)^2}\phi(t_k) + \lim_{n\to\infty} \sum_{k=0}^{n-1} \pi\epsilon\sqrt{1 + \dot{\phi}(t_k)^2}\epsilon\dot{\phi}(t_k).$$

We would like to ignore the second term. Since ϕ is continuously differentiable, $\dot{\phi}$ is a continuous function on the closed interval $[a, b]$ and hence it is bounded. That means that $\pi\sqrt{1 + \dot{\phi}(t_k)^2}\dot{\phi}(t_k) \leq C$ for some constant C. So

expression inside the limit in the second term is bounded above by

$$\sum_{k=0}^{n-1} C\epsilon^2 = \sum_{k=0}^{n-1} C(b-a)^2/n^2 = C(b-a)^2/n$$

which goes to zero in the limit as $n \to \infty$. Hence we can ignore this term.

The first term is actually just what we would get if we wrote out the definition of the Riemann integral of the function

$$2\pi\phi(t)\sqrt{1+\dot{\phi}(t)^2}$$

using the sequence of decompositions we began with. This proves the lemma. □

Theorem 38. *Assume that there is a $\mathcal{C}^2$-function $\phi\colon [a,b] \to \mathbf{R}$ which gives a surface of revolution of minimal area subject to the boundary conditions $\phi(a) = A$, $\phi(b) = B$. Then ϕ has the form*

$$\phi(x) = C\cosh((x-D)/C)$$

for some constants C, D to be determined by the boundary conditions.

This curve is called a catenary curve and that's why its surface of revolution is called a *catenoid*. We'll meet the catenary curve again in Section 2.4.3 below.

Proof. We are seeking to minimise the functional

$$F(\phi) = \int_{-1}^{1} \phi(t)\sqrt{1+\dot{\phi}(t)^2}dt$$

i.e.

$$L(p,q,r) = q\sqrt{1+r^2}$$

Clearly, L is independent of p, so we have Beltrami's identity

$$L - \dot{\phi}\frac{\partial L}{\partial \dot{\phi}} = C$$

for some constant C. Since $\frac{\partial L}{\partial \dot{\phi}} = \phi\dot{\phi}/\sqrt{1+\dot{\phi}^2}$, this becomes

$$\phi\sqrt{1+\dot{\phi}^2} - \frac{\dot{\phi}^2\phi}{\sqrt{1+\dot{\phi}^2}} = C.$$

Multiplying by $\sqrt{1+\dot{\phi}^2}$:

$$(1+\dot{\phi}^2)\phi - \dot{\phi}^2\phi = C\sqrt{1+\dot{\phi}^2}$$

or

$$\phi = C\sqrt{1+\dot{\phi}^2}.$$

This rearranges to give

$$\dot{\phi} = \sqrt{\frac{\phi^2}{C^2} - 1}.$$

Now

$$\begin{aligned} t + D' \equiv \int dt &= \int \frac{\dot{\phi}}{\sqrt{\frac{\phi^2}{C^2} - 1}} dt \\ &= C \int \frac{d(\phi/C)}{\sqrt{\frac{\phi^2}{C^2} - 1}} \\ &= C \log \left(2 \left(\frac{\phi}{C} + \sqrt{\frac{\phi^2}{C^2} - 1} \right) \right) \\ &= C \log 2 + C \cosh^{-1} \left(\frac{\phi}{C} \right) \end{aligned}$$

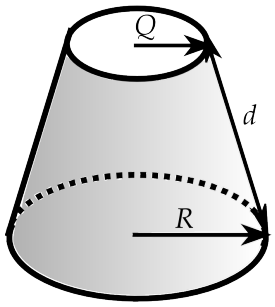
Therefore, setting $D = C \log 2 - D'$, we have

$$\phi(t) = C \cosh \left(\frac{t - D}{C} \right)$$

as required. □

Problem 39. *Suppose that $a = -1$, $b = 1$. Show that the equal height case $A = B$ leads us to $\phi(t) = C\cosh(t/C)$. For what values of A does this have a solution? Given the interpretation of a minimal surface of revolution as a soap-film, how do you think we should interpret this last answer?*

2.3.2 Area of a frustum



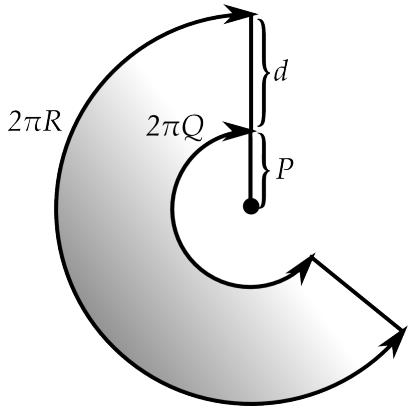
See `http://youtu.be/OEeJeBE1vg0`. A conical frustum is the surface of a truncated cone. Let Q and R denote the radii of the circles at the top and

bottom of a frustum (assume $Q < R$). In the proof of Lemma 37 we used the following formula for the surface area of a frustum:

$$\pi d(Q + R)$$

where d is the distance between the top and bottom circles as measured on the surface of the frustum. We will demonstrate this formula here, though the demonstration begins with an unjustified piece of geometric intuition that I won't bother to explain.

The point I will not justify is that you can cut the frustum and lay it out flat. The result is a planar region: a segment of an annulus. You could demonstrate this by writing down an explicit isometric (distance-preserving) map between the cut frustum and the planar region, but I won't do that. Given this fact, we will compute the area of the planar region.



Let F be the fraction of the annulus which is subtended by the cut frustum and let P denote the radius of the inner boundary of the annulus. We know that the distance between the two boundary components of the annulus is d. We also know that the length of the inner arc is $2\pi Q$ and the length of the outer arc is $2\pi R$. In other words,

$$2\pi PF = 2\pi Q$$
$$2\pi(P + d)F = 2\pi R.$$

This gives

$$F = \frac{(R - Q)}{d}$$

and

$$P = \frac{Qd}{R - Q}.$$

Therefore the area of the annulus, being the difference of the area of two discs, is

$$\pi\left(P + d\right)^2 - \pi P^2 = \pi d(d + 2P)$$

and the area of the segment is a fraction F of this, giving the area of the frustum as

$$\begin{aligned}\pi(R-Q)(d+2P) &= \pi(R-Q)\left(d+2\frac{Qd}{R-Q}\right)\\ &= \pi d(R-Q+2Q)\\ &= \pi d(Q+R)\end{aligned}$$

as claimed.

2.3.3 The brachistochrone

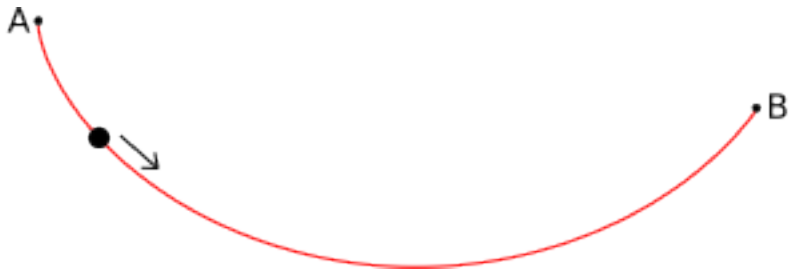


Figure 2.3: The brachistochrone curve is the optimal shape of wire for a frictionless bead moving under gravity to slide down between two points in order to minimise the time taken.

From the Greek $\beta\rho\alpha\chi\iota\varsigma\tau o\sigma + \chi\rho o\nu o\varsigma$ meaning "shortest" + "time", the brachistochrone problem is the problem of finding the shape of a wire joining two points p and q such that a bead running on this wire under gravity (no friction) gets from p to q in the shortest time (assuming it starts off at rest at p).

Let's translate p to the origin and suppose that q is at (x_0, y_0), with $x_0 > 0$ (so that p and q are not directly above and below one another) and with $y_0 < 0$ (so that the bead will really fall from p to q). The kinetic energy of a bead with speed v is $\frac{1}{2}mv^2$ and the gravitational potential energy of a bead at (x, y) is mgy. Therefore

$$\tfrac{1}{2}mv^2 + mgy$$

is constant throughout the trajectory. At p the bead is at rest and $y = 0$ so this whole expression vanishes identically. This means that we can solve for v:

$$v = \sqrt{-2gy}$$

We assume that the wire is the graph of some function $y(x)$. Moreover, we'll assume that this function is invertible, that is we can express x as a function of y and $dx/dy = (dy/dx)^{-1}$.

The time taken for the bead to move along the wire to q is

$$T = \int_p^q dt = \int_p^q \frac{ds}{\dot{s}}$$

where $s(t)$ is the distance travelled, so $\dot{s} = v$. Therefore

$$\begin{aligned} T &= \int_p^q \frac{ds}{v} \\ &= \int_p^q \frac{\sqrt{dx^2 + dy^2}}{\sqrt{-2gy}} \\ &= \int_p^q \frac{\sqrt{1 + \frac{dy}{dx}^2}}{\sqrt{-2gy}} dx \end{aligned}$$

and we have a calculus of variations problem on our hands: minimise T as a functional of the curve $y(x)$ (under the assumption that $y(0) = 0$ and $y(x_0) = y_0$). The Lagrangian for our problem is $L(p, q, r) = \frac{\sqrt{1+r^2}}{\sqrt{-2gq}}$.

The Lagrangian is independent of its first variable and so we have Beltrami's identity

$$\frac{\sqrt{1 + (y')^2}}{\sqrt{-2gy}} - y' \frac{\partial}{\partial y'} \left(\frac{\sqrt{1 + (y')^2}}{\sqrt{-2gy}} \right) = C$$

for some constant C. This gives

$$\frac{1}{\sqrt{-2gy}\sqrt{1 + (y')^2}} (1 + (y')^2 - (y')^2) = C$$

or

$$y' = \sqrt{\frac{1}{-2C^2 gy} - 1}.$$

Writing this in terms of $x(y)$, we get

$$\begin{aligned} \frac{dx}{dy} &= \frac{1}{\sqrt{\frac{1}{-2C^2 gy} - 1}} \\ &= \frac{\sqrt{-y}}{\sqrt{A + y}}, \end{aligned}$$

where $A = 1/2C^2 g$. We can now integrate this by substituting $y = -A \sin^2(\theta)$ and using the condition $x(0) = 0$ to get

$$x = A \sin^{-1} \sqrt{\frac{-y}{A}} - \sqrt{-y}\sqrt{A + y}.$$

Remark 40. *From Figure 2.3 it should be clear that our assumption (that x is expressible as a function of y) is a poor one. Of course, the end result is reassuringly multivalued (it's got a $\sin^{-1}$ and a $\sqrt{\cdot}$). So although we made an illegal assumption, the answer is actually correct. This is the beauty of the Euler-Lagrange equations: their* locality *- from a* global *assumption about a function (that it minimises a functional) you derive an equation for the* local *behaviour of that extremal function. In the same way it's kind of magical that light travels in such a way as to minimise the time taken during travel, by only obeying local laws of physics (without knowing beforehand what its endpoint is). If we were feeling careful we could make this brachistochrone argument more rigorously, but for the purposes of this course we'll just be satisfied that we got the right answer.*

2.4 Constrained problems

We have already considered constrained optimisation in the finite-dimensional setting. Exactly the same idea works in infinite-dimensional functional optimisation problems, but it is easier not have to think in terms of tangencies of infinite-dimensional hypersurfaces! I will explain how the algorithm works and give some examples.

2.4.1 The algorithm

The algorithm is the same as usual. You have a functional defined on some space of functions $\phi(t)$. This functional is given by an integral

$$F(\phi) = \int_a^b L(t, \phi, \dot{\phi})dt$$

and you want to find its extrema. However, you also want to impose a constraint on the functions ϕ. Maybe you want to fix the total arc-length of the graph of ϕ,

$$\int_a^b \sqrt{1 + \dot{\phi}^2}dt$$

or the area underneath the graph

$$\int_a^b \phi(t)dt.$$

Notice that these constraints are given by integrals; let's consider the general constraint:

$$\int_a^b M(t, \phi, \dot{\phi})dt = C.$$

We impose this by introducing a Lagrange multiplier $\lambda \in \mathbf{R}$. The full functional we want to consider is

$$F(\phi) = \int_a^b \left(L\left(t, \phi, \dot{\phi}\right) + \lambda \left(M\left(t, \phi, \dot{\phi}\right) - \frac{C}{b-a} \right) \right) dt.$$

Varying with respect to λ gives

$$\frac{\partial F}{\partial \lambda} = \int_a^b M\left(t, \phi, \dot{\phi}\right) dt - C = 0.$$

Varying with respect to ϕ gives the usual Euler-Lagrange equation but where L is replaced by $L + \lambda(M - C)$:

$$\frac{\partial}{\partial \phi}\left(L + \lambda(M-C)\right) = \frac{d}{dt}\left(\frac{\partial}{\partial \dot{\phi}}\left(L + \lambda(M-C)\right) \right).$$

2.4.2 The isoperimetric problem

For an interesting historical survey of this problem and approaches to proving it, see this article by V. Blåsjö in the American Mathematical Monthly.

The isoperimetric problem is to find a curve in the plane with a fixed length K which bounds an area A which is maximal amongst all areas bounded by plane curves of length K. The answer is, of course, a circle of radius $K/2\pi$. We will prove the following:

Theorem 41. *If there exists a 2π-periodic (i.e. closed) $\mathcal{C}^2$-curve $\gamma\colon \mathbf{R} \to \mathbf{R}^2$ of length K which maximises the area it bounds amongst all closed $\mathcal{C}^2$-curves of length K, then it is a circle.*

Note that, as usual, we do not prove existence of a maximiser! We will simply show that the only critical point of a suitable functional on the space of curves is the circle. In the problem sheets we will use Fourier analysis to prove the stronger statement that the circle is a maximiser.

Proof of Theorem 41. If C is a curve parametrised by a 2π-periodic $\mathcal{C}^2$-function $\gamma\colon \mathbf{R} \to \mathbf{R}^2$ and B is the bounded component of $\mathbf{R}^2 \backslash C$ then we seek to maximise the area integral $\int_B dxdy$ subject to the condition that $\int_0^{2\pi} \sqrt{\dot{\gamma}_1^2 + \dot{\gamma}_2^2} dt = K$.

Of course, this integral doesn't look quite right yet: we're used to integrating along a curve, not over an area. By Green's theorem:

$$\int_B dxdy = \int_B d(xdy) = \int_C xdy = \int_0^{2\pi} \gamma_1(t)\dot{\gamma}_2(t)dt$$

so our functional becomes

$$F(\gamma) = \int_0^{2\pi} \left(\gamma_1(t)\dot{\gamma}_2(t) - \lambda\left(\sqrt{\dot{\gamma}_1^2 + \dot{\gamma}_2^2} - \frac{K}{2\pi} \right) \right) dt$$

where λ is a Lagrange multiplier. Let's write x and y rather than γ_1 and γ_2. The Lagrangian is

$$L(x, y, \dot{x}, \dot{y}) = x\dot{y} - \lambda\left(\sqrt{\dot{x}^2 + \dot{y}^2} - \frac{K}{2\pi}\right)$$

and the Euler-Lagrange equations are

$$\dot{y} = -\frac{d}{dt}\left(\frac{\lambda\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}}\right)$$
$$0 = \frac{d}{dt}\left(x - \frac{\lambda\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}}\right)$$
$$K = \int_0^{2\pi} \sqrt{\dot{x}^2 + \dot{y}^2}dt$$

(differentiating $F(\gamma)$ in the γ_1, γ_2 and λ-directions respectively).

Now we do something cunning, which we also did when we dealt with straight lines. If we are given a path $[a, b] \to \mathbf{R}^2$, we can reparametrise that path to get a new path $[a, b] \to \mathbf{R}^2$. This way we can make a particle travel along it at any speed $v(t)$ we want, provided that

$$\int_a^b v(t)dt = K$$

and we certainly don't change the length of the path (the distance the particle travels). For the rigorous justification of this, see Section 2.1.3. In particular, we can parametrise *proportionally to arc-length* and define our new time coordinate $s(t) = \frac{2\pi}{K}\int_0^t \sqrt{\dot{\gamma}_1^2 + \dot{\gamma}_2^2}dt$. This satisfies

$$\frac{d}{ds}\int_0^s \sqrt{\dot{\gamma}_1(\sigma)^2 + \dot{\gamma}_2(\sigma)^2}d\sigma = \frac{K}{2\pi}$$

in other words

$$\sqrt{\left(\frac{d\gamma_1}{ds}\right)^2 + \left(\frac{d\gamma_2}{ds}\right)^2} = \frac{K}{2\pi}$$

or

$$\sqrt{\dot{x}^2 + \dot{y}^2} = \frac{K}{2\pi}.$$

The Euler-Lagrange equations simplify to

$$\dot{y} = -\frac{2\pi\lambda}{K}\ddot{x}$$
$$\dot{x} = \frac{2\pi\lambda}{K}\ddot{y}$$

(note that the other equation is automatically satisfied). We have

$$\frac{d^3y}{ds^3} = -\frac{K^2}{4\pi^2\lambda^2}\frac{dy}{ds}$$

so $\dot{y}$ is a harmonic oscillator. This implies that the solution is in fact a circle (could be centred anywhere and parametrised starting from any angle). □

It seems slightly devious that, halfway through a Lagrange-multiplier proof, we reparametrised so we could ignore the Lagrange-multiplier equation. The question I asked myself when I started off writing down this proof was: Why don't we restrict attention to arc-length parametrisations from the very beginning? If you try it, you'll get the wrong answer. The reason is that the space of paths parametrised by arc-length is not a nice flat space. If you add an arbitrary perturbation $\gamma+\epsilon$, you completely ruin the condition that the curve is parametrised by arc-length. For the same reason, we have to impose the length-fixing constraint with a Lagrange multiplier instead of by starting off with the space of all curves of a fixed length.

So why were we allowed to reparametrise then? Because we knew that, if a solution of the constrained problem existed, we could reparametrise it afterwards and what we obtained would still be an extremum of the constrained problem. So we might as well restrict ourselves to curves parametrised by arc-length, but we really need to be using the constrained Euler-Lagrange equation.

2.4.3 The catenary

Coming from the Latin word "catena" meaning chain, the catenary is the curve which follows a hanging chain. More formally, if (a, A) and (b, B) are points in a vertical plane and there is an infinitely thin chain of uniform density ρ (mass per unit length) whose endpoints are fixed at these points, hanging freely under gravity and having fixed length K then the chain fits a curve called the catenary, given by the graph of a function $\phi(t)$, $t \in [a, b]$ which is the solution to a variational problem.

The variational problem is this: minimise the total potential energy of the chain subject to the condition that its length equals K. The potential energy of a mass at height y is mgy so each "infinitesimal element" of our chain $(x(t), y(t)) = (t, f(t))$ has:

- length equal to $\sqrt{dx^2 + dy^2} = \sqrt{1 + \dot{f}^2}dt$,
- mass equal to $\rho\sqrt{1 + \dot{f}^2}dt$,
- potential energy equal to $\rho g f\sqrt{1 + \dot{f}^2}dt$.

Integrating this over the whole interval $[a, b]$ gives the total potential energy

$$\rho g \int_a^b f(t)\sqrt{1 + \dot{f}(t)^2}dt$$

and the total length is

$$K = \int_a^b \sqrt{1 + \dot{f}(t)^2} dt.$$

The functional for the constrained variational problem is therefore

$$\rho g \int_a^b f(t)\sqrt{1 + \dot{f}(t)^2} dt + \lambda \left(\int_a^b \sqrt{1 + \dot{f}(t)^2} dt - K \right)$$

where λ is a Lagrange multiplier.

Problem 42. *Find the constrained Euler-Lagrange equations for this problem and show that the solution is a catenary curve*

2.5 Several dimensions

The problems we have considered so far have all concerned optimising functions of one variable. The resulting Euler-Lagrange equation is a second-order ordinary differential equation. Things become even more interesting if we consider variational problems with functions of several variables. The theory is almost identical except that ordinary derivatives become partial derivatives and the equations become harder!

2.5.1 Euler-Lagrange equation in two variables

The generalisation to many variables will hopefully be clear. We consider a functional which takes functions $\phi(x, y)$ of two variables and outputs an integral

$$\int \int L(x, y, \phi, \partial_x \phi \, \partial_y \phi) dx dy$$

where $L(p, q, r, s, t)$ is a function of five variables. The usual Euler-Lagrange argument applies but we need to use the several-variable version of the chain rule. This results in the equation

$$\frac{\partial L}{\partial q} = \frac{\partial}{\partial x}\left(\frac{\partial L}{\partial s}\right) + \frac{\partial}{\partial y}\left(\frac{\partial L}{\partial t}\right)$$

where all the derivatives of L are evaluated at

$$(p, q, r, s, t) = (x, y, \phi(x, y), (\partial_x \phi)(x, y), (\partial_y \phi)(x, y)).$$

In order to reflect this, the equation is usually written

$$\boxed{\frac{\partial L}{\partial \phi} = \frac{\partial}{\partial x}\left(\frac{\partial L}{\partial \phi_x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial L}{\partial \phi_y}\right)}.$$

where we are using the notation $\phi_x = \partial\phi/\partial x$, $\phi_y = \partial\phi/\partial y$ simply to minimise the number of ∂s floating around.

2.5.2 Laplace's equation

Consider a plane region B bounded by a closed curve C and let $f\colon C \to \mathbf{R}$ be a function. We seek a $\mathcal{C}^2$-function $\varphi\colon B \to \mathbf{R}$ satisfying $\varphi|_C = f$ and minimising the following integral over all such functions:

$$F(\varphi) = \int_B \left((\partial_x\varphi)^2 + (\partial_y\varphi)^2\right) dxdy = \int_B |\nabla\varphi|^2 dxdy.$$

Since

$$\frac{\partial L}{\partial \varphi} = 0$$

the Euler-Lagrange equation is

$$0 = \frac{\partial}{\partial x}\left(2\frac{\partial\varphi}{\partial x}\right) + \frac{\partial}{\partial y}\left(2\frac{\partial\varphi}{\partial y}\right) = 2\Delta\varphi$$

where Δ denotes the Laplace operator

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}.$$

The extremals therefore satisfy $\Delta\varphi = 0$ (Laplace's equation) and are *harmonic functions*. We will meet Laplace's equation again later as the steady-temperature case of the heat equation: temperature distributions evolve over time via the heat equation

$$\frac{\partial\varphi}{\partial t} = \Delta\varphi$$

and, after a long time when the temperature reaches a steady state so that $\partial\varphi/\partial t = 0$, we achieve a steady-temperature distribution satisfying Laplace's equation. You can see from our variational characterisation that the heat equation is acting to minimise the total gradient of the temperature distribution (as defined by the functional $F(\varphi)$). This fits well with our physical intuition.

2.5.3 Minimal surface equation

Let $C \subset \mathbf{R}^2$ be a closed curve in the plane bounding a region B and let $f\colon C \to \mathbf{R}$ be a given function. We would like to find an extension φ of f to the whole of B which minimises the surface area of $\mathrm{Graph}(\varphi)$. This is precisely what a soap-film would do if you dipped the graph of f (considered as a wire frame) into a suitably soapy liquid mixture.

Lemma 43. *The surface area of* $\mathrm{Graph}(\varphi)$ *is*

$$\int_B \sqrt{1 + \left(\frac{\partial\varphi}{\partial x}\right)^2 + \left(\frac{\partial\varphi}{\partial y}\right)^2}\, dxdy.$$

Proof. I'll only give a sketch proof. Take a tangent plane to the graph of φ at a point on the graph. Two of the vectors in this tangent plane are $(1, 0, \partial_x f)$ and $(0, 1, \partial_y f)$. The area of the parallelogram which they span is equal to the magnitude of their cross-product

$$(1, 0, \partial_x f) \times (0, 1, \partial_y f) = (-\partial_x f, -\partial_y f, 1)$$

and this magnitude is precisely the integrand. Now approximate the surface by such parallelograms (living over squares of edge-length ϵ in $\mathbf{R}^2$) and take better and better approximations ($\epsilon \to 0$). This yields the integral. □

Theorem 44. *If φ is a $\mathcal{C}^2$-function with $\varphi|_C = f$ which minimises the surface area of its graph amongst all such functions then φ satisfies*

$$\frac{\partial^2 \varphi}{\partial x^2}\left(1 + \left(\frac{\partial \varphi}{\partial y}\right)^2\right) + \frac{\partial^2 \varphi}{\partial y^2}\left(1 + \left(\frac{\partial \varphi}{\partial x}\right)^2\right) = 2\frac{\partial \varphi}{\partial x}\frac{\partial \varphi}{\partial y}\frac{\partial^2 \varphi}{\partial x \partial y}.$$

The proof is consigned to one of the problem sets: I could not deprive you of the pleasure of this calculation.

Part II

Partial differential equations

Chapter 3

Some general theory

3.1 The definition

A partial differential equation (PDE) for a function $\phi(x_1, \ldots, x_n)$ of n variables is a relation between the values of ϕ and (finitely many of) its partial derivatives (to arbitrary orders) at each point. That is, the relation never relates the partial derivatives or values of ϕ at different points - that would be a *difference equation*, and would probably be harder to solve. We will write $F(\phi)(x, y) = 0$ for this relation at each point (x, y). We should probably write something like

$$F\left(x, y, \phi(x, y), (\partial_x \phi)(x, y), (\partial_y \phi)(x, y), (\partial_x^2 \phi)(x, y), \ldots\right) = 0$$

where the dots indicate a finite list, but that would be cumbersome.

For example,

$$F(\phi)(x, y) = \frac{\partial \phi}{\partial x}(x, y) + \phi(x, y)\frac{\partial \phi}{\partial y}(x, y) - xy = 0$$

or

$$\left(\frac{\partial \phi}{\partial x}\right)^2 + \left(\frac{\partial \phi}{\partial y}\right)^2 = 1$$

or

$$\frac{\partial \phi}{\partial t} = \frac{\partial^2 \phi}{\partial x^2}$$

are PDEs, where we have already started to omit the argument from each term. On the other hand,

$$\partial_x \phi(x, y) = \partial_y \phi(x + 1, y - 1)$$

and

$$\partial_x \phi + \partial_x^2 \phi + \partial_x^3 \phi + \cdots = 0$$

are not PDEs (one is a difference equation, one involves infinitely many derivatives).

The highest order of derivatives which appears in the equation is called the *order* of the equation. We will first consider first-order equations, then second-order equations. But here are some more definitions:

Definition 45.

- *A PDE $F(\phi) = 0$ is called linear if*

$$F(\lambda_1\phi_1 + \lambda_2\phi_2) = \lambda_1 F(\phi_1) + \lambda_2 F(\phi_2)$$

for all $\lambda_1, \lambda_2 \in \mathbf{R}$. In practical terms, this means that $F(\phi)$ is a sum of terms of the form

$$A\partial_1 \cdots \partial_k \phi$$

where A depends only on the variables $x_1, \ldots, x_n$ and $\partial_1 \ldots \partial_k$ is some string of partial derivatives. We also allow inhomogeneous linear equations, $F(\phi) = R$ where F is linear and R is a function of $x_1, \ldots, x_n$.

- *If moreover, each coefficient A is just a constant, we say the equation has constant coefficients.*

- *A PDE is called quasilinear if, in each term involving highest (say kth) order partial derivatives $\partial_1 \cdots \partial_k \phi$, the coefficient doesn't involve any other kth derivatives. So*

$$\frac{\partial\phi}{\partial x}\frac{\partial\phi}{\partial y} + x\frac{\partial\phi}{\partial y} = 1$$

is not quasilinear because it is first order but there is a term involving a product of first derivatives. The Burgers inviscid fluid equation

$$\phi\frac{\partial\phi}{\partial x} + \frac{\partial\phi}{\partial t} = 0$$

is quasilinear.

- *If a PDE is not quasilinear then it is called* fully nonlinear*. For example the eikonal equation*

$$\left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial y}\right)^2 = 1$$

important in geometric optics or the Monge-Ampère equation

$$\frac{\partial^2\phi}{\partial x^2}\frac{\partial^2\phi}{\partial y^2} - \left(\frac{\partial^2\phi}{\partial x\partial y}\right)^2$$

important in geometry.

Most of the PDEs we are going to talk about in this course in are linear with the exception of a few quasilinear and fully nonlinear first order equations (like the eikonal equation). We will start by talking about first-order equations because there is a very well-developed theory for them. Then we will talk about some specific second-order equations. I will take a moment to discuss the second-order equations we're going to study.

3.2 Quasilinear second order equations in two variables

The general quasilinear second order equation in two variables (x, y) is

$$A\frac{\partial^2\phi}{\partial x^2} + B\frac{\partial^2\phi}{\partial x\partial y} + C\frac{\partial^2\phi}{\partial y^2} = R$$

where A, B, C, R are arbitrary functions of $x, y, \phi, \partial_x\phi$ and $\partial_y\phi$.

Definition 46. *Define $\Delta = B^2 - 4AC$. This is a function of $x, y, \phi, \partial_x\phi, \partial_y\phi$. However, sometimes it happens that this function is always positive, always negative or always zero. For instance, if the equation is linear and has constant coefficients then Δ is just a number.*

- *If $\Delta < 0$ we say that the equation is elliptic,*
- *if $\Delta = 0$ we say that the equation is parabolic,*
- *and if $\Delta > 0$ we say that the equation is hyperbolic.*

This is only a classification for linear, constant coefficient equations for which $\Delta > / = / < 0$ makes sense. More generally, it is a guideline *- for example, we expect solutions to quasilinear elliptic equations to share certain properties with solutions to linear elliptic equations.*

Here are some examples of linear equations with constant coefficients.

Example 47. *Laplace's equation*

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = 0$$

has $A = C = 1$, $B = 0$ so $\Delta = -4 < 0$ which means it is an elliptic equation.

Example 48. *The heat equation*

$$\frac{\partial\phi}{\partial t} = \frac{\partial^2\phi}{\partial x^2}$$

has $A = 1, B = 0, C = 0$. Therefore $\Delta = 0$ and the equation is parabolic.

Example 49. *The wave equation*

$$\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = \frac{\partial^2\phi}{\partial x^2}$$

has $A = 1/c^2, B = 0, C = -1$ so $\Delta = 1/c^2 > 0$ and the equation is hyperbolic.

This classification may seem like something quite formal, but the three types of equations really do display very different kinds of behaviour (and quasilinear equations of elliptic type really do behave very like linear elliptic equations).

Problem 50. *Show that the (quasilinear) minimal surface equation*

$$\frac{\partial^2\varphi}{\partial x^2}\left(1+\left(\frac{\partial\varphi}{\partial y}\right)^2\right)+\frac{\partial^2\varphi}{\partial y^2}\left(1+\left(\frac{\partial\varphi}{\partial x}\right)^2\right)=2\frac{\partial\varphi}{\partial x}\frac{\partial\varphi}{\partial y}\frac{\partial^2\varphi}{\partial x\partial y}$$

from Section 2.5.3 is elliptic.

3.3 Some basic tricks

Let's restrict our attention to linear, homogeneous equations in two variables with constant coefficients. We take first- and second-order equations and look at some basic tricks for solving them. These tricks amount to making clever changes of coordinates and will generalise to the *method of characteristics* for first-order equations. They will also make it clear why Δ is a natural quantity to consider for second order equations.

3.3.1 First order

Suppose we have a first-order equation we want to solve:

$$A\frac{\partial\phi}{\partial x}+B\frac{\partial\phi}{\partial y}+C\phi=0.$$

The expression $A\frac{\partial\phi}{\partial x}+B\frac{\partial\phi}{\partial y}$ is just the partial derivative of ϕ in the (A,B)-direction. Consider the family of all lines parallel to the vector (A,B), that is the lines

$$\{(x_0+tA,y_0+tB)\ :\ t\in\mathbf{R}\}$$

Suppose we have a solution ϕ and we restrict it to one of these lines:

$$u(t)=\phi(x_0+tA,y_0+tB)$$

Then we know

$$\frac{du}{dt}=A\frac{\partial\phi}{\partial x}+B\frac{\partial\phi}{\partial y}$$

so

$$\frac{du}{dt}(t)+Cu(t)=0.$$

This is an ordinary differential equation equivalent to

$$\frac{d}{dt}(e^{Ct}u)=0$$

and has solution

$$u(t)=Ke^{-Ct}$$

Suppose that $A \neq 0$. Let $x_0 = 0$. Then, as t and y_0 vary, the point $(tA, y_0 + tB)$ traces out the whole of $\mathbf{R}^2$. Indeed, the point (x, y) corresponds to $t = x/A$, $y_0 = y - xB/A$. Therefore

$$\phi(x, y) = K(y - xB/A)e^{-Cx/A}.$$

Note that we are allowing the 'constant of integration' K to be an arbitrary function of $y - xB/A = y_0$ because y_0 was fixed when we restricted to the line $(At, y_0 + tB)$. This makes sense because $y - xB/A$ is annihilated by the directional derivative

$$A\frac{\partial}{\partial x} + B\frac{\partial}{\partial y}.$$

This basic trick is the idea behind the method of characteristics: using the original equation, find a system of curves (in this case straight lines) along which the PDE reduces to an ODE which we can solve.

3.3.2 Second order

Consider the equation

$$A\frac{\partial^2\phi}{\partial x^2} + B\frac{\partial^2\phi}{\partial x \partial y} + C\frac{\partial^2\phi}{\partial y^2} = 0 \tag{3.1}$$

and suppose that $\Delta = B^2 - 4AC > 0$, i.e. the equation is hyperbolic. Then the quadratic polynomial

$$At^2 + Bt + C = 0 \tag{3.2}$$

has two distinct real roots

$$t_\pm = \frac{-B \pm \sqrt{\Delta}}{2A}$$

so that $At^2 + Bt + C = A(t - t_-)(t - t_+)$. Define $s_\pm = y + xt_\pm$ so that

$$x = \frac{s_+ - s_-}{t_+ - t_-}, \qquad y = \frac{t_+ s_- - t_- s_+}{t_+ - t_-}.$$

Then

$$\begin{aligned}
\frac{\partial}{\partial s_+} &= \frac{\partial x}{\partial s_+}\frac{\partial}{\partial x} + \frac{\partial y}{\partial s_+}\frac{\partial}{\partial y} \\
&= \frac{1}{t_+ - t_-}\left(\frac{\partial}{\partial x} - t_-\frac{\partial}{\partial y}\right) \\
\frac{\partial}{\partial s_-} &= \frac{\partial x}{\partial s_-}\frac{\partial}{\partial x} + \frac{\partial y}{\partial s_-}\frac{\partial}{\partial y} \\
&= -\frac{1}{t_+ - t_-}\left(\frac{\partial}{\partial x} - t_+\frac{\partial}{\partial y}\right).
\end{aligned}$$

This means that

$$\begin{aligned}\frac{\partial}{\partial s_+}\frac{\partial}{\partial s_-} &= -\frac{1}{(t_+ - t_-)^2}\left(\frac{\partial}{\partial x} - t_-\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial x} - t_+\frac{\partial}{\partial y}\right) \\ &= -\frac{1}{A(t_+ - t_-)^2}\left(A\frac{\partial^2}{\partial x^2} + B\frac{\partial^2}{\partial x \partial y} + C\frac{\partial^2}{\partial y^2}\right)\end{aligned}$$

and hence, if ϕ solves (3.1),

$$\frac{\partial^2 \phi}{\partial s_+ \partial s_-} = 0.$$

The general solution to this equation is a sum of two completely arbitrary differentiable functions:

$$f(s_+) + g(s_-),$$

i.e.

$$f(y + xt_+) + g(y + xt_-).$$

If the original equation were instead parabolic then the quadratic polynomial (3.2) would have two identical real roots t_0 and the general solution would be $f(y+xt_0)+xg(y+xt_0)$. To see this, use the new coordinates $(u, v) = (y+t_0x, x)$ instead of $s_\pm$.

If the original equation were elliptic then (3.2) would have two non-real complex-conjugate roots z and z^*. The general solution is once again

$$f(y + xz) + g(y + xz^*)$$

but the functions f and g (which may be complex) have to be chosen so as to make their sum real. For instance,

$$f(t) = g(t) = t$$

would give

$$y + xz + y + xz^* = 2y + x\,\mathrm{Re}(z)$$

as the solution.

Examples

Example 51. *Suppose we wish to find the general solution to*

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial x \partial y} + -2\frac{\partial^2 \phi}{\partial y^2} = e^x.$$

The quadratic equation is $t^2+t-2 = (t-1)(t+2)$ *and therefore* $t_- = -2,\ t_+ = 1$. *Then* $s_- = y - 2x$ *and* $s_+ = y + x$ *and the equation becomes*

$$-9\frac{\partial^2 \phi}{\partial s_+ \partial s_-} = \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial x \partial y} + -2\frac{\partial^2 \phi}{\partial y^2} = e^x.$$

Note that $x = (s_+ - s_-)/3$ *so our equation is*

$$\frac{\partial^2 \phi}{\partial s_+ \partial s_-} = \frac{-e^{s_+ - s_-}}{9}$$

Integrating with respect to s_+ *and* s_- *gives*

$$\phi = e^{s_+ - s_-} + F(s_+) + G(s_-)$$

or

$$\phi(x, y) = e^x + F(y + x) + G(y - 2x).$$

Now let us see how to solve for F and G when given initial conditions $\phi(x, 0) = M(x)$ and $\partial\phi/\partial t(x, 0) = N(x)$.

Example 52. *Say we require* $\phi(x, 0) = x^2$ *and* $\partial\phi/\partial t(x, 0) = x^3$. *Then*

$$x^2 = e^x + F(x) + G(-2x)$$
$$x^3 = F'(x) + G'(-2x)$$

Differentiating the first of these gives

$$2x = e^x + F'(x) - 2G'(-2x)$$

and now we have two simultaneous equations for F' *and* G'. *These give*

$$3F'(x) + e^x = 2(x^3 + x)$$

or

$$F(x) = \frac{x^4}{6} + \frac{x^2}{3} - \frac{e^x}{3}$$

Plugging back into the very first equation gives

$$G(-2x) = x^2 - e^x - F(x)$$

or

$$G(-2x) = \frac{2x^2}{3} - \frac{x^4}{6} - \frac{2e^x}{3}.$$

Substituting $z = -2x$ *gives*

$$G(x) = \frac{z^2}{6} - \frac{z^4}{96} - \frac{2e^{-z/2}}{3}.$$

Hence

$$\phi(x, y) = e^x + \frac{(x + y)^4}{6} + \frac{(x + y)^2}{3} - \frac{e^{x+y}}{3} + \frac{(y - 2x)^2}{6} - \frac{(y - 2x)^4}{96} - \frac{2e^{-(y-2x)/2}}{3}.$$

Chapter 4

First order PDE

As usual, I will only deal with two-variable equations, mainly for notational convenience.

4.1 Linear case

We have already seen how to solve first order linear homogeneous PDE with constant coefficients in Section 3.3.1. Let's just quickly extend this to the inhomogeneous case. The equation we want to solve is

$$A\frac{\partial \phi}{\partial x} + B\frac{\partial \phi}{\partial y} + C\phi = D(x, y)$$

The trick we employed before works again: we restrict attention to the line $(At, y_0 + Bt)$ along which the equation for $u(t) = \phi(At, y_0 + Bt)$ reduces to an ODE

$$\frac{du}{dt} + Cu = D(At, y_0 + Bt)$$

or

$$\frac{d}{dt}(e^{Ct}u) = e^{Ct}D(At, y_0 + Bt)$$

and the solution consists of a particular integral

$$u(t) = e^{-Ct}\int_0^t e^{Cs}D(As, y_0 + Bs)ds$$

plus a complementary function

$$K(y_0)e^{-Ct}$$

with a 'constant' depending on y_0, and we can substitute $t = x/A$, $y_0 = y - Bx/A$ as before to solve the original problem:

$$\phi(x,y) = e^{-Cx/A}\int_0^{x/A} e^{Cs}D(As, y - Bx/A + Bs)ds + K(y - Bx/A)e^{-Cx/A}.$$

4.2 Linear case, varying coefficients

Now we allow A, B, C to depend on (x, y) too (but not yet on ϕ). The equation we want to solve is

$$A(x,y)\frac{\partial\phi}{\partial x} + B(x,y)\frac{\partial\phi}{\partial y} + C(x,y)\phi = D(x,y).$$

The operator $A(x,y)\frac{\partial\phi}{\partial x} + B(x,y)\frac{\partial\phi}{\partial y}$ has an obvious interpretation as a directional derivative, but for the *varying vector field* $v(x,y) = (A(x,y), B(x,y))$. We need a replacement for the family of straight lines (u, y_0) in the previous case.

Definition 53. *An* integral curve *for the vector field $v(x,y)$ is a curve $\gamma\colon \mathbf{R} \to \mathbf{R}^2$ such that*

$$\dot\gamma(t) = v(\gamma(t)),$$

in other words, a curve whose tangent vector is everywhere given by v.

Example 54. *Suppose that $v(x,y) = (A, B)$, where A and B are constant. Then the lines $(x_0, y_0) + t(A, B)$ are integral curves for v.*

Example 55. *Suppose $v(x,y) = (-y, x)$. Then the family of circles of radius r, $\gamma(t) = (r\cos(t), r\sin(t))$, are integral curves for v. Indeed,*

$$\dot\gamma(t) = (-r\sin(t), r\cos(t)) = v(\gamma(t)).$$

We'll sort out the problem of actually *finding the integral curves* momentarily. Let's assume for the moment that we have found them. If $\gamma(t)$ is a parametrisation of an integral curve then

$$\frac{d}{dt}\phi(\gamma(t)) = d_{\gamma(t)}\phi(\dot\gamma(t)) = d_{\gamma(t)}\phi(v)$$

which is just the directional derivative of ϕ in the v-direction at $\gamma(t)$, in other words

$$A(\gamma(t))\frac{\partial\phi}{\partial x} + B(\gamma(t))\frac{\partial\phi}{\partial y}.$$

Therefore, setting $f(t) = \phi(\gamma(t))$, we have

$$\frac{df}{dt}(t) = D(\gamma(t)) - C(\gamma(t))f(t)$$

which is an ordinary differential equation for f.

So the only difficulty is in working out γ. We have

$$\frac{d\gamma_1}{dt} = A(\gamma_1(t), \gamma_2(t))$$
$$\frac{d\gamma_2}{dt} = B(\gamma_1(t), \gamma_2(t))$$

which is a system of two ODEs for two unknowns γ_1, γ_2.

Therefore in order to solve the PDE, we have to solve a pair of coupled ODEs for γ and then a further ODE to find f along γ. This latter ODE introduces a 'constant of integration', but this 'constant' is allowed to vary as you move from one integral curve γ to another.

Definition 56 (Characteristics)**.** *The curves $\gamma(t)$ are called the* characteristics *of our equation. Note that when we solve the system of ODEs for γ we get two constants of integration and hence a two-parameter family of curves. One of these parameters can always be fixed because it will correspond to reparametrising the characteristics.*

Let's do an example.

Example 57. *Consider*

$$\frac{\partial \phi}{\partial x} + 2x\frac{\partial \phi}{\partial y} = 1.$$

We have $A(x, y) = 1$, $B(x, y) = 2x$ and hence

$$\frac{d\gamma_1}{dt} = 1$$
$$\frac{d\gamma_2}{dt} = 2\gamma_1$$

which has solution $\gamma_1(t) = t + M$ and $\gamma_2(t) = (t + M)^2 + N$; in other words, the integral curves are parabolae. Let us fix $M = 0$ by translating the time parameter.

Now for each N define $f(t) = \phi(t, t^2 + N)$ and we need to solve $df/dt = 1$. This gives $f(t) = t + K(N)$ or, given that $x = t$ and $y - x^2 = N$,

$$\phi(x, y) = x + K(y - x^2)$$

which is the general solution of our problem.

Example 58. *Consider*

$$y\frac{\partial \phi}{\partial x} + x\frac{\partial \phi}{\partial y} + xy\phi = 0.$$

We have $A(x,y)=y$, $B(x,y)=x$ *and hence*

$$\frac{d\gamma_1}{dt}=\gamma_2$$
$$\frac{d\gamma_2}{dt}=\gamma_1$$

or

$$\frac{d(\gamma_1-\gamma_2)}{dt}=-(\gamma_1-\gamma_2)$$
$$\frac{d(\gamma_1+\gamma_2)}{dt}=\gamma_1+\gamma_2$$

which has solution $\gamma_1-\gamma_2=Me^{-t}$, $\gamma_1+\gamma_2=Ne^{t}$ *or*

$$\gamma(t)=\left(\frac{1}{2}\left(Ne^{t}+Me^{-t}\right),\frac{1}{2}\left(Ne^{t}-Me^{-t}\right)\right)$$

which are the hyperbolae $x^2-y^2=MN$. *Let us fix* $M=1$ *and solve for* $f(t)=\phi\left(\frac{1}{2}\left(Ne^{t}+Me^{-t}\right),\frac{1}{2}\left(Ne^{t}-Me^{-t}\right)\right)$.

$$\begin{aligned}\frac{df}{dt}&=-xyf\\&=-\frac{1}{4}\left(Ne^{t}+e^{-t}\right)\left(Ne^{t}-e^{-t}\right)f(t)\\&=-\frac{1}{4}\left(N^2e^{2t}-e^{-2t}\right)f(t)\end{aligned}$$

which has solution

$$f(t)=K(N)\exp\left(-\frac{1}{8}\left(N^2e^{2t}+e^{-2t}\right)\right)$$

or, using the fact that $Ne^{t}=x+y$, $e^{-t}=x-y$,

$$\begin{aligned}\phi(x,y)&=K(x^2-y^2)\exp\left(-\frac{1}{8}\left((x+y)^2+(x-y)^2\right)\right)\\&=K(x^2-y^2)\exp\left(-\frac{1}{4}\left(x^2+y^2\right)\right)\end{aligned}$$

4.3 Quasilinear case

See `http://youtu.be/43i-A6lt_MI`.

Now we move on to quasilinear equations

$$A(x,y,\phi)\frac{\partial\phi}{\partial x}+B(x,y,\phi)\frac{\partial\phi}{\partial y}+C(x,y,\phi)=0. \tag{4.1}$$

This is slightly harder to deal with: instead of using a system of two coupled ODEs to find the characteristics and then solving the final ODE to work out f along the characteristics, we end up with a system of three coupled ODEs for γ_1, γ_2, f all at the same time and we have no obvious way of separating out the three.

4.3.1 Characteristic vector field

Suppose that ϕ is a solution to the equation and that $(x(t), y(t), f(t))$ is the restriction of this solution to a curve $(x(t), y(t))$. The system of ODE is

$$\frac{dx}{dt} = A(x(t), y(t), f(t)) \tag{4.2}$$

$$\frac{dy}{dt} = B(x(t), y(t), f(t)) \tag{4.3}$$

$$\frac{df}{dt} = -C(x(t), y(t), f(t)). \tag{4.4}$$

You can see that we are unlikely to be able to solve for x and y first because f occurs in all three equations.

Suppose this system of equations is satisfied by some family of curves $(x_s(t), y_s(t), z_s(t))$, that is: for each s the curve

$$(x_s(t), y_s(t), z_s(t))$$

satisfies (4.2)-(4.4). This family of curves traces out some surface in (x, y, z)-space. Suppose that this surface is the graph of a function $z = \phi(x, y)$. A simple application of the chain rule gives:

$$\frac{d\phi(x_s(t), y_s(t))}{dt} = \frac{dx_s}{dt}\frac{\partial \phi}{\partial x} + \frac{dy_s}{dt}\frac{\partial \phi}{\partial y}$$

and when we substitute in the right-hand sides of (4.2)-(4.4) we see that ϕ satisfies (4.1) identically!

This means that if a function $\phi(x, y)$ has the property that its graph

$$\{z = \phi(x, y)\}$$

coincides with the surface traced out by

$$(s, t) \mapsto (x_s(t), y_s(t), z_s(t))$$

then ϕ is a solution to Equation (4.1).

Conversely, if ϕ is a solution to (4.1) then $(x(t), y(t), \phi(x(t), y(t)))$, i.e. $z(t) = \phi(x(t), y(t))$, is a solution curve to the system (4.2)-(4.4) of ordinary differential

equations, since

$$\begin{aligned}\frac{dz}{dt} &= \frac{dx}{dt}\frac{\partial\phi}{\partial x} + \frac{dy}{dt}\frac{\partial\phi}{\partial y} \\ &= A\frac{\partial\phi}{\partial x} + B\frac{\partial\phi}{\partial y} \\ &= -C\end{aligned}$$

as required.

Definition 59. *The vector field $(A, B, -C)$ in (x, y, z)-space is called the* characteristic vector field *of the quasilinear equation* (4.1)*. The integral curves of $(A, B, -C)$ are called characteristic curves and the projections of the characteristic curves to the xy-plane are called the* characteristic projections*.*

In this language, what we have proved is that (a) the graph of a solution to (4.1) is a union of characteristic curves, and that (b) a union of characteristic curves traces out a surface which, when it is a graph, is the graph of a solution to (4.1).

4.3.2 Cauchy data

We know that the graph of our solution is a union of characteristic curves. We have to pick an 'initial condition' for our PDE. This will take the form of a curve in (x, y, z)-space.

Cauchy Problem. *Given a curve $\chi\colon \mathbf{R} \to \mathbf{R}^2$ in (x, y)-space and a function $f\colon \chi \to \mathbf{R}$ defined over that curve, find a solution ϕ to*

$$A\phi_x + B\phi_y + C = 0$$

such that $\phi|_\chi = f$. This initial condition is called the Cauchy data and the solution is called its development. The curve χ is called the Cauchy hypersurface.

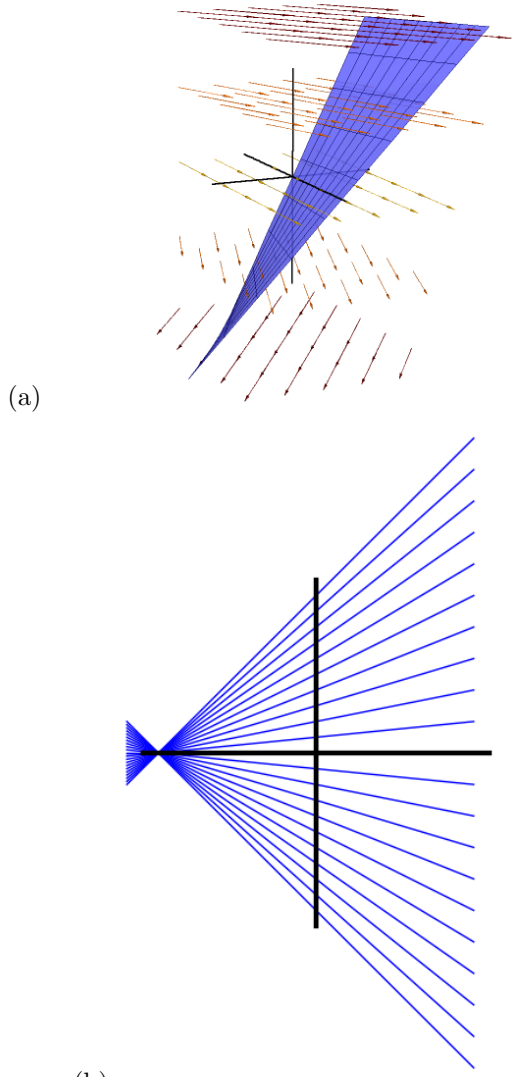
Geometrically, solving the Cauchy problem means finding the characteristic curves which pass through the curve

$$s \mapsto (\chi_1(s), \chi_2(s), f(\chi(s)))$$

in (x, y, z)-space. The union of these characteristic curves is called a *solution surface* and is parametrised by the coordinates (s, t) (where s is a coordinate on the Cauchy hypersurface and t is a coordinate along the characteristic curve).

4.3.3 An example: Burgers's equation

It's high time we solved an example.



(a)

(b)

Figure 4.1: (a) The solution surface to $\frac{\partial \phi}{\partial t} + \phi \frac{\partial \phi}{\partial y} = 0$ with the Cauchy data $\phi(0, y) = y$, plotted with the characteristic vector field. This surface is a union of straight lines which are characteristic curves. (b) The characteristic projections of this solution. You can see that they begin to cross at $(-1, 0)$.

Example 60. *Take the equation*

$$\frac{\partial \phi}{\partial x} + \phi \frac{\partial \phi}{\partial y} = 0$$

which has $A(x,y,z) = 1$, $B(x,y,z) = z$, $C(x,y,z) = 0$. Therefore the characteristic vector field is

$$\begin{aligned}\dot{x} &= 1\\ \dot{y} &= z\\ \dot{z} &= 0\end{aligned}$$

which has integral curves

$$\begin{aligned}x &= t + D\\ y &= Et + F\\ z &= E\end{aligned}$$

for constants D, E, F. Without loss of generality we can take $D = 0$ because we can reparametrise t so that $x(0) = 0$. These characteristic curves are straight lines, pointing horizontally but with a slope in the xy-plane which increases as z increases.

Let's choose some Cauchy data. We'll take as Cauchy hypersurface the straight line $\chi = \{(0, y) \in \mathbf{R}^2\}$, which is parametrised by $s \mapsto (0, s)$. We'll take two different sets of Cauchy data: $\phi(0, s) = f(s)$ with (a) $f(s) = s$, (b) $f(s) = s^2$. This means that in case (a) we are looking for the characteristic curves which pass through the points $(0, s, s)$ in (x, y, z)-space and in case (b) we are looking for the characteristic curves which pass through the points $(0, s, s^2)$ in (x, y, z)-space.

We first deal with case (a). The point $(0, s, s)$ intersects the characteristic curve $(t, Et + F, E)$ at $t = 0$ if $E = F = s$. In other words, the Cauchy data pick out the characteristic curves

$$t \mapsto (t, st + s, s).$$

This means that $(s, t) \mapsto (t, st+s, s)$ is a parametrisation of our solution surface by the coordinates (s, t).

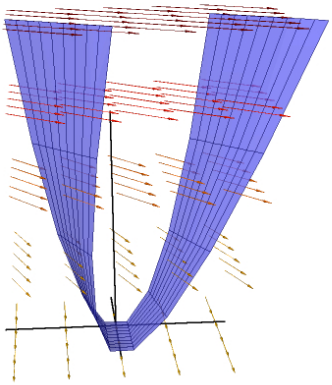
For case (b), the point $(0, s, s^2)$ intersects the characteristic curve $(t, Et+F, E)$ at $t = 0$ if $F = s$, $E = s^2$. In other words, the Cauchy data pick out the characteristic curves

$$t \mapsto (t, s^2t + s, s^2).$$

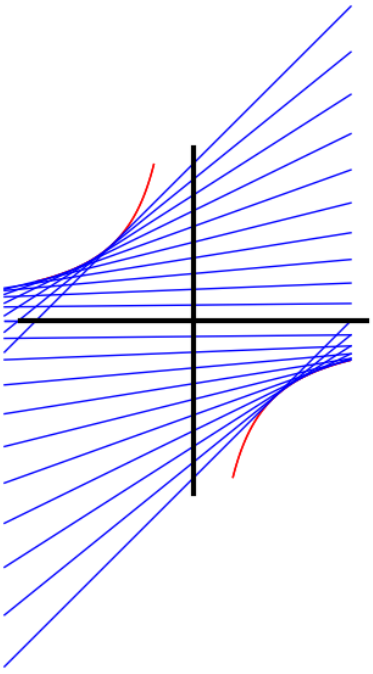
This means that $(s, t) \mapsto (t, s^2t + s, s^2)$ is a parametrisation of our solution surface by the coordinates (s, t).

4.3.4 Caustics

The problem with this method is that is produces solution surfaces which are parametrised by (s, t) instead of by (x, y)! It is quite possible that the surface $S \subset \mathbf{R}^3$ traced out as (s, t) vary is not the graph of any function $z = \phi(x, y)$.



(a)



(b)

Figure 4.2: (a) The solution surface to $\frac{\partial \phi}{\partial t} + \phi \frac{\partial \phi}{\partial y} = 0$ with Cauchy data $\phi(0, y) = y^2$, plotted with the characteristic vector field. You can see the graph starting to bend over and become double-valued. (b) The characteristic projections of this solution. You can see that they are all tangent to the red curve $1 + 4xy = 0$ and that they are starting to cross one another near that curve.

Indeed, we can see in Figures 4.1 and 4.2 that the Cauchy development is usually not a graph. Let's see what goes wrong in the previous example if we try and express ϕ as a function of (x, y).

In case (a), the solution surface is parametrised by

$$(s, t) \mapsto (t, st + s, s)$$

so $y = z(x + 1)$ and $z = y/(x + 1)$. When $x = -1$ this doesn't make sense. This manifests itself in our picture: all the characteristic projections cross at the point $(-1, 0)$.

In case (b), the solution surface is parametrised by

$$(s, t) \mapsto (t, s^2t + s, s^2)$$

so $y = zx + \sqrt{z}$. If $u = \sqrt{z}$ then $xu^2 + u - y = 0$ and so

$$z = u^2 = \left(\frac{-1 \pm \sqrt{1 + 4xy}}{2x}\right)^2.$$

this doesn't make sense when $1 + 4xy < 0$ and, even when $1 + 4xy$, gives two possible values for $z = \phi(x, y)$. This manifests itself in our picture: all the characteristic projections live in the region $1 + 4xy \geq 0$ and they are all tangent to the curve $1 + 4xy = 0$ (in red in the figure). The solution surface simply doesn't extend as far as $1 + 4xy > 0$.

The red curve in Figure 4.2 is the *envelope* of the characteristic curves, in other words it is the curve to which all of the characteristics are somewhere tangent. In the case of the eikonal equation of geometric optics, where characteristics are light trajectories, this envelope shows up as a very bright curve called the *caustic*.

Remark 61. *The equation from our example is known as Burgers's equation (without viscosity) and is usually written*

$$\frac{\partial \phi}{\partial t} + \phi \frac{\partial \phi}{\partial x} = 0.$$

The interpretation is that the x-axis is filled with a fluid and the velocity of fluid particles at x at time t is $\phi(x, t)$. The phenomenon of a single-valued solution becoming multiple-valued is known as shock-wave formation and corresponds to the fluid "catching up with itself" so that fluid particles with different velocities coexist at the same (x, t)-coordinate.

We want to understand the caustic more generally. It is actually quite easy to calculate the caustic straight from the parametrisation of the solution surface by (s, t).

Definition 62. *The* caustic *of a Cauchy development*

$$(s,t) \mapsto (u(s,t), v(s,t), w(s,t))$$

is the set of critical values of the map

$$\pi\colon \mathbf{R}^2 \to \mathbf{R}^2, \qquad \pi(s,t) = (u(s,t), v(s,t)),$$

in other words, the projection under π of the set of points $p \in \mathbf{R}^2$ where

$$\det \begin{pmatrix} \partial_s u & \partial_t u \\ \partial_s v & \partial_t v \end{pmatrix} (p) = 0.$$

I will quickly explain why the caustic is relevant, then we will proceed to compute it in some examples. Remember that the graph of a function $\phi\colon \mathbf{R}^2 \to \mathbf{R}$ is just the surface

$$(x,y) \mapsto (x, y, \phi(x,y)).$$

Our solution surfaces have the form

$$(s,t) \mapsto (u(s,t), v(s,t), w(s,t))$$

so that x and y are given implicitly in terms of s and t. Therefore it is not clear if s, t or w can be given in terms of x and y. If w cannot be expressed in terms of x and y then the solution surface fails to be the graph of a function: the surface "folds over itself". In other words, the projection

$$\pi(s,t) = (u(s,t), v(s,t))$$

which forgets the value of w is not one-to-one. Along the fold, the solutions surface admits vertical tangencies, that is vectors tangent to the solution surface which point vertically upward in the w-direction. So we can detect the bad behaviour of our solution by looking for vertical tangencies to our solution surface. It turns out that this is precisely the problem of computing the determinant in the definition of the caustic.

Lemma 63. *If the solution surface $(s,t) \mapsto (u(s,t), v(s,t), w(s,t))$ has a vertical tangency at some point (s_0, t_0) then $(u(s,t), v(s,t))$ is in the caustic.*

Proof. Suppose without loss of generality that $s_0 = t_0 = 0$. Consider the curves $s \mapsto (u(s,0), v(s,0), w(s,0))$ and $t \mapsto (u(0,t), v(0,t), w(0,t))$. These are curves on the solution surface passing through the point which has a vertical tangency. Their velocity vectors:

$$X = (\partial_s u(0,0), \partial_s v(0,0), \partial_s w(0,0)) \text{ and } Y = (\partial_t u(0,0), \partial_t v(0,0), \partial_t w(0,0))$$

span the tangent space of the solution surface at that point. In other words, all tangent vectors are linear combinations $AX + BY$ of these two. If one of the tangent directions $AX + BY$ is vertical then when it is projected

into the (u, v)-plane it vanishes, and so the corresponding linear combination vanishes: $A\pi(X)+B\pi(Y) = 0$ (where π is the projection $\pi(u, v, w) = (u, v)$). The vectors $\pi(X)$ and $\pi(Y)$ are therefore linearly dependent. The vectors $\pi(X)$ and $\pi(Y)$ are the rows of the matrix

$$\begin{pmatrix} \partial_s u & \partial_t u \\ \partial_s v & \partial_t v \end{pmatrix}(p)$$

appearing in the definition of the caustic (where p is the point $(0, 0)$). The determinant of a matrix vanishes if and only if its rows are linearly dependent, which proves the lemma. □

Example 64. *In our earlier example, case (a), we had*

$$\pi(s, t) = (u(s, t), v(s, t)) = (t, st + s)$$

and so the determinant of the matrix of partial derivatives is

$$\det \begin{pmatrix} 0 & 1 \\ t+1 & s \end{pmatrix} = -t - 1$$

which vanishes when $t = -1$. When $t = -1$, we have $(u(s, -1), v(s, -1)) = (-1, 0)$ so the caustic is a single point: the point we discovered earlier as the intersection of all the characteristic projections.

In case (b) we had $(u(s, t), v(s, t)) = (t, s^2t + s)$ and so the determinant of the matrix of partial derivatives is

$$\det \begin{pmatrix} 0 & 1 \\ 2st+1 & s^2 \end{pmatrix} = -2st - 1$$

which vanishes when $2st = -1$. This means that (x, y) is on the caustic if it has the form $(u(-1/2t, t), v(-1/2t, t)) = (t, -1/4t)$ so the caustic is the curve $\{4xy + 1 = 0\}$.

Example 65. *Consider the PDE*

$$-\sin\phi\frac{\partial\phi}{\partial x} + \cos\phi\frac{\partial\phi}{\partial y} = 1$$

for which the characteristic vector field is

$$(-\sin(z), \cos(z), 1).$$

The solution to the characteristic system of ODE is

$$(x(t), y(t), z(t)) = (x_0 + \cos t, y_0 + \sin t, t)$$

(reparametrising so that $z(0) = 0$) and these characteristic curves are helices, spiraling upward. Suppose that the initial condition is $\phi(x, 0) = 0$ so that the

Cauchy hypersurface is $\chi(s) = (s, 0)$ *and the Cauchy data is* $\phi(s, 0) = 0$. *The characteristic curve* $(x_0 + \cos t, y_0 + \sin t, t)$ *intersects* $(s, 0, 0)$ *at* $t = 0$ *if* $y_0 = 0$ *and* $x_0 = s - 1$, *so the solution surface is parametrised by*

$$(s, t) \mapsto (s - 1 + \cos t, \sin t, t)$$

Clearly the solution is $\phi(x, y) = \sin^{-1}(y)$ *but this is not single-valued, indeed there are infinitely many* z *such that* $\sin z = y$. *Moreover, the solution goes no further than* $y = 1$ *- this corresponds to the fact that helices start overlapping when* $t = \pi/2$ *(see Figure 4.3: the characteristics are segments of circles with centre on the* x*-axis).*

The matrix of first derivatives of the projection of the solution surface to the (x, y)*-plane is*

$$\begin{pmatrix} 1 & -\sin t \\ 0 & \cos t \end{pmatrix}$$

which has determinant $\cos t$. *This vanishes precisely along* $t = \pi/2$ *as suspected, so the caustic is the line* $y = \sin(\pi/2) = 1$.

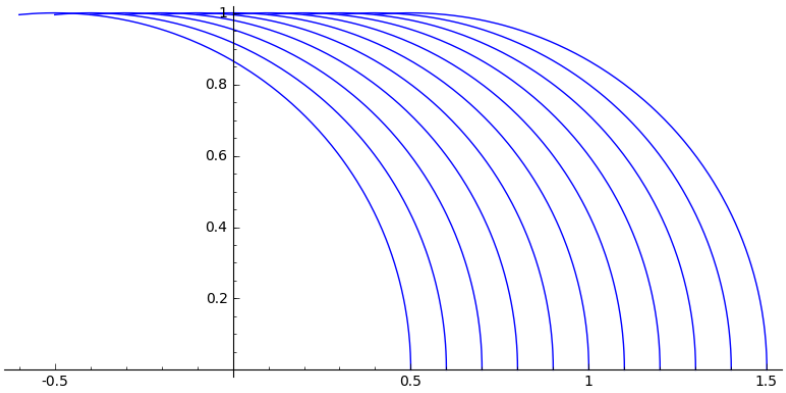


Figure 4.3: The characteristics of $-\sin\phi\frac{\partial\phi}{\partial x} + \cos\phi\frac{\partial\phi}{\partial y} = 1$ are segments of circle. We see they start to overlap near $y = 1$.

4.4 Fully nonlinear case

The method works for the fully nonlinear case but the system of ODEs becomes yet more complicated.

Suppose that $G(x, y, u, p, q)$ is a function of five variables and that $\phi\colon \mathbf{R}^2 \to \mathbf{R}$ is a function satisfying

$$G\left(x, y, \phi(x, y), \partial_x\phi(x, y), \partial_y\phi(x, y)\right) = 0$$

for all $(x, y) \in \mathbf{R}^2$. Now let $\gamma(t) = (x(t), y(t))$ be a curve and restrict ϕ to γ to obtain a function $u(t)$ as usual. Let $G(t)$ denote

$$G(x(t), y(t), u(t), p(t), q(t))$$

where $p(t) = \partial_x \phi(x(t), y(t))$, $q(t) = \partial_y \phi(x(t), y(t))$. We have

$$\frac{dG}{dt} = \frac{\partial G}{\partial x}\dot{x} + \frac{\partial G}{\partial y}\dot{y} + \frac{\partial G}{\partial u}\dot{u} + \frac{\partial G}{\partial p}\dot{p} + \frac{\partial G}{\partial q}\dot{q}$$

and since $\dot{u} = \dot{x}\frac{\partial \phi}{\partial x} + \dot{y}\frac{\partial \phi}{\partial y} = \dot{x}p + \dot{y}q$ this becomes

$$\begin{aligned}\frac{dG}{dt} = \dot{x}\left(\frac{\partial G}{\partial x} + \frac{\partial G}{\partial u}p\right) + \dot{y}\left(\frac{\partial G}{\partial y} + \frac{\partial G}{\partial u}q\right) + \\ + \frac{\partial G}{\partial p}\dot{p} + \frac{\partial G}{\partial q}\dot{q}.\end{aligned}$$

This suggests a system of five coupled ODEs for the five quantities $(x(t), y(t), u(t), p(t), q(t))$:

$$\begin{aligned}\frac{dx}{dt} &= \frac{\partial G}{\partial p} & \dot{p} &= -\left(\frac{\partial G}{\partial x} + p\frac{\partial G}{\partial u}\right) \\ \frac{dy}{dt} &= \frac{\partial G}{\partial q} & \dot{q} &= -\left(\frac{\partial G}{\partial y} + q\frac{\partial G}{\partial u}\right) \\ & & \dot{u} &= \frac{\partial G}{\partial p}p + \frac{\partial G}{\partial q}q\end{aligned}$$

As usual, if we integrate this system of ODEs we will obtain a curve. Taking a one-parameter family of these integral curves gives a surface in (x, y, u, p, q)-space and when we project to (x, y, u)-space we obtain a surface which, wherever it is a graph, is the graph of a solution $u = \phi(x, y)$.

Example 66. *Let us consider the eikonal equation (in units where the speed of light is 1)*

$$\left(\frac{\partial \phi}{\partial x}\right)^2 + \left(\frac{\partial \phi}{\partial y}\right)^2 = 1$$

(so $G(x, y, u, p, q) = p^2 + q^2$) which describes the time ϕ taken by light emitted normally by some curve $C \subset \mathbf{R}^2$ to each a point $(x, y) \in \mathbf{R}^2$. To see that this description is valid, let's consider the corresponding system of ODEs:

$$\begin{aligned}\frac{dx}{dt} &= 2p & \dot{p} &= 0 \\ \frac{dy}{dt} &= 2q & \dot{q} &= 0 \\ & & \dot{u} &= 2(p^2 + q^2) = 2\end{aligned}$$

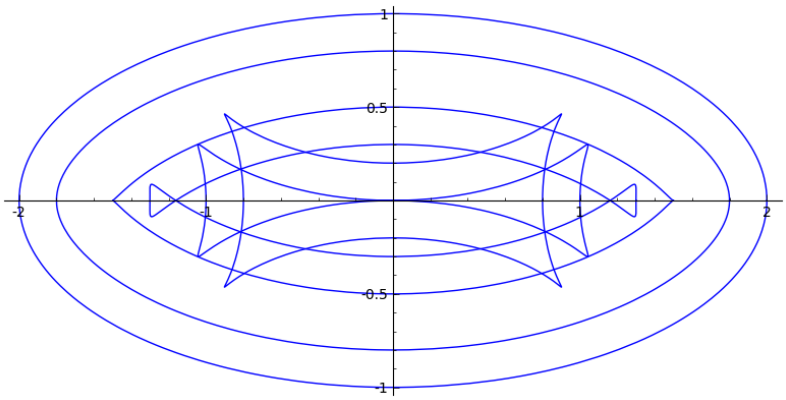


Figure 4.4: The equidistants of an ellipse form singularities as characteristics cross.

Starting from the curve C and choose the initial condition $\phi|_C = 0$. We see that this determines p and q along C, namely (p, q) must be (plus or minus) the unit normal vector to the curve C because $p^2 + q^2 = 1$ (giving unit length) and because $(p, q) = \left(\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}\right)$ and by our choice of initial condition the directional derivative of ϕ along C vanishes, so (p, q) is normal to C. Now along the integral curves of the ODE, p and q do not change and hence x and y follow the normal line (with speed 2) and the solution to $\dot{u} = 2$ is just $u(t) = 2t$. If one followed the normal with speed 1, we would get $u(t) = t$.

This is precisely the statement that the solution of the eikonal equation is the time taken by light to reach (x, y) from C in a normal direction. The characteristics are straight lines normal to C. Solutions of the eikonal equation can be very beautiful. Figure 4.4 shows some of the singularities developed by level sets of ϕ corresponding to the initial condition $C = \left\{\frac{x^2}{4} + y^2 = 1\right\}$ (that is, the equidistants of an ellipse).

Chapter 5

The heat equation and Laplace's equation

5.1 The heat equation

We want to study the evolution of a temperature distribution which is not in a steady state. To simplify matters we restrict to one-dimensional distributions - these are then represented by functions $\phi\colon [0,L]\times[0,\infty)\to\mathbf{R}$ where the first coordinate is spatial (x) and the second is time (t) and the equation governing their evolution is the heat equation, a parabolic second order PDE:

$$\frac{\partial\phi}{\partial t}=\frac{\partial^2\phi}{\partial x^2}.$$

5.1.1 Boundary conditions

We need to impose boundary conditions in the x and t-directions. The t-boundary conditions are called initial conditions (for obvious reasons!). We will always impose an initial condition like

$$\phi(x,0)=F(x)$$

for some function F, and then seek a solution on $[0,L]\times[0,\infty)$.

There are various options for the boundary conditions in the x-direction:

- You can specify $\phi(0,t)$ and $\phi(L,t)$. This is called the Dirichlet problem.
- Instead of specifying ϕ along one of these boundaries you can specify $\frac{\partial\phi}{\partial x}(0,t)$ or $\frac{\partial\phi}{\partial x}(L,t)$ (in other words the directional derivative of ϕ in a direction normal to the boundary). This is called a Neumann boundary

condition. You could also impose $\frac{\partial \phi}{\partial t}(x, 0)$, i.e. a Neumann initial condition! But let's not do that.

- Mixtures of the two are perfectly OK.

I was always confused by Neumann boundary conditions. What's the point? With Dirichlet it's clear what it means - like with a soap film filling in a fixed boundary shape. Here is a physical situation (heat) in which both are clearly relevant.

- **Dirichlet:** The boundary points of our interval are kept at fixed temperatures T_1, T_2, maybe by immersing the region in some vast bath of refridgerating liquid. Since our interval is one-dimensional then you've got to imagine immersing it in a one-dimensional bath, so that only the boundary is being refridgerated...
- **Neumann:** The boundary of our interval is completely insulated from the outside world. Insulation means that heat cannot flow into or out of the boundary. Since heat travels down a temperature gradient, this means that the directional derivative in the normal direction to the boundary must vanish.

5.1.2 Strategy

We now want to solve the heat equation with given boundary and initial conditions. Fourier solved this problem in his treatise on heat in 1822. His idea was:

- Find lots of solutions to the heat equation. Make sure they fit the boundary conditions. Maybe they don't fit your initial condition, but don't worry.
- Take linear superpositions of (possibly infinitely many of) these solutions in such a way as to fit the resulting function to the initial condition.

This works because the heat equation is linear, so if $\phi_1, \phi_2, \ldots$ are solutions and $A_1, A_2, \ldots$ are real numbers then $\sum_{k=1}^{\infty} A_k \phi_k$ is also a solution (provided that the partial sums converge in an appropriate way (uniformly) so that derivatives commute with $\sum_{k=1}^{\infty}$).

5.1.3 Separation of variables

We find lots of solutions by using a method called *separation of variables*: we seek solutions of the form

$$\phi(x, y) = X(x)T(t)$$

where $X(x)$ is a function of x alone and $T(t)$ is a function of t alone. In this case the heat equation becomes

$$XT' = X''T.$$

If X and T solve $X'' = \lambda X$ and $T' = \lambda T$ then certainly $XT' - X''T = \lambda XT - \lambda XT = 0$ so the problem would be solved. I didn't just pull these equations out of a hat, the argument usually goes like this: divide through by XT to get

$$\frac{X''}{X} = \frac{T'}{T} = \lambda$$

for some constant λ (has to be constant because X''/X depends only on X and T'/T depends only on T). Multiplying up gives the two equations I mentioned. This has the disadvantage of dividing by XY which could potentially vanish. In future we will not be made queasy by this division and, from now on, that's how we'll derive equations for separation of variables.

Solving for T, we get either

$$T(t) = Ke^{\lambda t} \text{ or } K$$

with K constant, according to whether $\lambda \neq 0$ or $\lambda = 0$ respectively. The $\lambda = 0$ case gives $X'' = 0$ and hence $X = Ax + B$. Since these have no time-dependence (because T is constant) these are called steady solutions.

The $\lambda > 0$ solutions have T increasing exponentially in time, which seems very unphysical! A hot plate, left to cool, doesn't suddenly become hotter. We will see shortly how our boundary conditions allow us to ignore such solutions.

We can write down the general separated solution by solving the ODEs above:

$$\phi_\lambda(x,t) = \begin{cases} \left(A\cos(x\sqrt{-\lambda}) + B\sin(x\sqrt{-\lambda})\right)e^{\lambda t} & \text{if } \lambda < 0 \\ (Ax + B) & \text{if } \lambda = 0 \\ \left(A\cosh(x\sqrt{-\lambda}) + B\sinh(x\sqrt{-\lambda})\right)e^{\lambda t} & \text{if } \lambda > 0 \end{cases}.$$

5.1.4 Fitting to boundary conditions

We fit the separated solutions to the boundary conditions. We specialise to Dirichlet conditions for simplicity. In the question sheets there will be some with Neumann conditions; the method is the same.

The Dirichlet conditions we impose are $\phi(0,t) = S_0$, $\phi(L,t) = S_1$. The steady solution

$$\phi_0(x,t) = \left(\frac{S_1 - S_0}{L}\right)x + S_0$$

obviously fits these conditions. By setting $\theta(x,t) = \phi(x,t) - \phi_0(x,t)$ get a new solution θ to the heat equation (because the heat equation is linear) satisfying

the new Dirichlet conditions $\theta(0,t) = 0 = \theta(L,t)$. In other words, we can assume without loss of generality that $S_0 = S_1 = 0$.

If $\lambda > 0$ then $X(x) = A\cosh(x\sqrt{\lambda}) + B\sinh(x\sqrt{\lambda})$ and so the condition $X(0) = 0$ means that $A = 0$. Also, the condition $X(L) = 0$ gives $B\sinh(L\sqrt{\lambda}) = 0$, but sinh only vanishes at 0, hence $B = 0$ also. Therefore we can ignore separated solutions with $\lambda > 0$ (as we had hoped) because we cannot fit them to our boundary conditions.

If $\lambda = -p^2 < 0$ then $X(x) = A\cos(px) + B\sin(px)$ and so the condition $X(0) = 0$ means that $A = 0$. Also, the condition $X(L) = 0$ gives $B\sin(pL) = 0$, and now sin vanishes at $n\pi$, $n \in \{1, 2, 3, ...\}$. This means we can have nontrivial solutions (i.e. with $B \neq 0$) whenever $pL = n\pi$. Thus we have separated solutions

$$\phi_n = e^{-n^2 t}\sin(n\pi x/L)$$

for $n = 1, 2, 3, ...$

5.1.5 Fitting to initial condition

Finally we want to take a (possibly infinite) linear combination of separated solutions ϕ_n to fit to the initial condition $\phi(x,0) = F(x)$. We are allowing ϕ_0 too. So the aim is to write

$$F(x) = \frac{S_1 - S_0}{L}x + S_0 + \sum_{n=1}^{\infty} A_n \sin(n\pi x/L)$$

in other words, we want to expand $F(x) - \frac{S_1 - S_0}{L}x - S_0$ as a (half-range) Fourier sine series. In other words, extend $F\colon [0,L] \to \mathbf{R}$ to an odd function $\tilde{F}\colon [-L,L] \to \mathbf{R}$:

$$\tilde{F}(x) = \begin{cases} F(x) & \text{if } x \in [0,L] \\ -F(-x) & \text{if } x \in [-L,0] \end{cases}$$

and take

$$A_n = \frac{1}{L}\int_{-L}^{L} \tilde{F}(x)\sin(n\pi x/L)dx.$$

No wonder Fourier invented Fourier series!

Example 67. *Suppose we want to solve $\partial\phi/\partial t = \partial^2\phi/\partial x^2$ for $x \in [0,\pi]$, subject to $\phi(x,0) = \cos x$, $\phi(0,t) = 1$, $\phi(\pi,t) = -1$. We start by setting*

$$\phi_0(x,t) = \frac{S_1 - S_0}{\pi}x + S_0 = 1 - 2x/\pi$$

and defining $\theta(x,t) = \phi(x,t) - \phi_0(x,t)$. Then θ solves the Dirichlet conditions

$$\theta(x,0) = \cos x - 1 + 2x/\pi,\ \theta(0,t) = 0 = \theta(\pi,t).$$

The solution to this problem is

$$\theta(x,t) = \sum_{n=1}^{\infty} A_n e^{-n^2 t} \sin(nx)$$

where A_n is the nth half-range sinusoidal Fourier coefficient of

$$F(x) = \cos x - 1 + 2x/\pi$$

which is

$$\frac{2}{\pi}\frac{((-1)^n+1)}{n(n^2-1)}$$

so that

$$\phi(x,t) = 1 - \frac{2x}{\pi} + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{((-1)^n+1)}{n(n^2-1)}e^{-n^2 t}\sin(nx).$$

5.1.6 The smoothing effect of parabolicity

In the limit as $t \to \infty$ all of the sinusoidal oscillations decay exponentially and we are left with the steady solution $A_0 x + B_0$. The Fourier modes with large n decay at a faster rate $e^{-n^2 t}$. This makes physical sense: heat moves down a temperature gradient to even out the temperature distribution. The more gradient there is, the faster the heat moves. The higher n is, the wigglier the Fourier mode, the steeper the gradient, the faster the heat moves to smooth out the temperature distribution. This behaviour is typical of parabolic PDEs.

5.2 Laplace's equation

5.2.1 Harmonic functions

The Laplace operator for functions of n variables $x_1, \ldots, x_n$ is

$$\Delta \sum_{k=1}^{n} \frac{\partial^2}{\partial x_k^2}.$$

Recall (we only wrote it out for $n = 2$) that this arises as the Euler-Lagrange equation for functions defined on a region $U \subset \mathbf{R}^n$ with specified boundary values, extremising the functional

$$\int_U \left(\sum_{k=1}^{n}\left(\frac{\partial \phi}{\partial x_k}\right)^2\right) dx_1 \cdots dx_n$$

known as the *Dirichlet energy*. Functions which are annihilated by Δ are called harmonic functions. The problem of finding a harmonic function on a given

region with specified boundary values is called Dirichlet's problem. The equation $\Delta\phi = 0$ satisfied by a harmonic function is called Laplace's equation; it is a linear elliptic second-order equation with constant coefficients.

Example 68. *Consider harmonic functions on the interval* $[0, 1]$*. Since these satisfy* $d^2f/dx^2 = 0$ *they are just functions of the form* $f(x) = ax + b$*, whose graphs are straight line segments from* $(0, b)$ *to* $(1, a + b)$*.*

This example displays behaviour typical of solutions to elliptic problems.

- After fixing the boundary conditions (in this case fixing b and $a+b$) there is only a very limited set of solutions (in this case a unique one).
- The solution is heavily dependent on the boundary conditions: if you change the boundary data locally (i.e. at just one of the boundary points), the solution is altered everywhere (i.e. arbitrarily close to the other boundary). This is in contrast with the wave equation: we will see that it takes time for changes in initial conditions to propagate (at the 'speed of light').
- The solution displays the following *mean value property*: $\frac{1}{2}\left(f(x) + f(y)\right) = f\left(\frac{x+y}{2}\right)$. Note that the solution is actually linear, but in higher dimensions it's only a weaker mean value property which holds, namely that the value of a harmonic function at the centre of a disc $D \subset U$ is equal to the integral of the function around the boundary of the disc (rescaled by the reciprocal of the circumference of the disc).

5.2.2 The maximum principle

We alluded above to the *mean value property*:

Theorem 69 (Mean value property of harmonic functions)**.** *If* $f\colon U \to \mathbf{R}$ *is a harmonic function on a domain* $U \subset \mathbf{R}^2$ *and* $D \subset U$ *is a disc centred at* x *of radius* r *then*

$$f(x) = \frac{\int_0^{2\pi} f(x + re^{i\theta})rd\theta}{\int_0^{2\pi} rd\theta}.$$

which is the analogue of the property

$$f\left(\frac{x+y}{2}\right) = \frac{1}{2}(f(x) + f(y))$$

for harmonic functions on an interval. We won't prove this, but note that it bears a marked (and non-coincidental) resemblance to Cauchy's integral formula in complex analysis! This formula has the consequence that the maxima and minima of a harmonic function are achieved along the boundary of its domain.

Corollary 70 (Maximum principle)**.** *If* $f\colon U \to \mathbf{R}$ *is a harmonic function on a domain* $U \subset \mathbf{R}^2$ *then the maximum and minimum of* f *are achieved along the*

boundary of U. If the maximum or minimum is also achieved in the interior of U then f is constant.

Proof. Suppose that the maximum is achieved somewhere (say x) in the interior and that the function is not constant. Then there exists a point y such that $f(y) < f(x)$. Let $U_1, \ldots, U_n$ be a sequence of discs such that x is at the centre of U_1, y is on the boundary of U_n and the boundary of U_k passes through the centre of U_{k-1} (we are making some topological assumptions about the domain U, like the fact that any two points in U can be connected by such a sequence of discs - that's a fairly mild condition).

Now let z_k be the point at the centre of U_k (so $x = z_1$), let $z_{n+1} = y$, and let r_k be the radius of U_k. Since $f(p) \leq f(x)$ for all $p \in U$ we have

$$\frac{\int_{\partial U_1} f(x + r_1 e^{i\theta}) r_1 d\theta}{\int_{\partial U_1} r_1 d\theta} \leq \frac{\int_{\partial U_1} f(x) r_1 d\theta}{\int_{\partial U_1} r_1 d\theta} = f(x)$$

with equality if and only if $f(p) = f(x)$ for all $p \in U_1$. By the mean value theorem, equality holds, so$f(z_1) = f(x)$. Applying the same argument again and again we eventually get to $f(y) = f(x)$, which is a contradiction. □

Corollary 71 (Uniqueness of solutions to Dirichlet problem)**.** *Suppose that $\phi_1, \phi_2 \colon U \to \mathbf{R}$ are harmonic functions on a domain $U \subset \mathbf{R}^2$ and $\phi_1|_{\partial U} = \phi_2|_{\partial U}$. Then $\phi_1 \equiv \phi_2$.*

Proof. Take $\theta = \phi_1 - \phi_2$. By linearity of the Laplace equation, $\Delta\theta = 0$. Moreover, $\theta|_{\partial U} \equiv 0$. By the maximum principle, the maximum and minimum of θ are achieved along ∂U, where it vanishes, hence it vanishes identically and $\phi_1 \equiv \phi_2$. □

5.2.3 Steady temperature distributions

Suppose that $U \subset \mathbf{R}^n$ is a region in n-dimensional Euclidean space (we're mostly interested in $n = 1, 2$) and that $\phi \colon \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$ is a time-dependent temperature distribution (also written $\phi(x, t)$, where $x \in U$ is a spatial variable and $t \in \mathbf{R}$ represents time). Then ϕ evolves according to the *heat equation*

$$\frac{\partial \phi}{\partial t} = \Delta \phi.$$

We see that harmonic functions correspond to steady temperature distributions, that is distributions which don't change over time. From the variational interpretation of Laplace's equation, these are the distributions which minimise the total temperature gradient (unsurprising, given that heat flows down a temperature gradient (from hot to cold) and tries to homogenise the temperature, subject to the boundary conditions).

5.2.4 Solution on a rectangular domain

We will now seek to solve the Laplace equation on a rectangular domain $U = [0, a] \times [0, b] \subset \mathbf{R}^2$. That is we need to find a harmonic function on the rectangle with arbitrary boundary data. Our strategy is the same as for the heat equation:

- Find lots of solutions to the Laplace equation. Maybe they don't fit your boundary values, but don't worry.
- Take linear superpositions of (possibly infinitely many of) these solutions in such a way as to fit the resulting function to the boundary data.

This works because the Laplace equation is linear, so if $\phi_1, \phi_2, \ldots$ are solutions and $A_1, A_2, \ldots$ are real numbers then $\sum_{k=1}^{\infty} A_k \phi_k$ is also a solution (provided that the partial sums converge in an appropriate way (uniformly) so that Δ commutes with $\sum_{k=1}^{\infty}$).

Separation of variables

We find lots of solutions by using a method called *separation of variables*: we seek solutions of the form

$$\phi(x, y) = X(x)Y(y)$$

where $X(x)$ is a function of x alone and $Y(y)$ is a function of y alone. Laplace's equation becomes

$$X''(x)Y(y) + X(x)Y''(y) = 0 \tag{5.1}$$

where X'' (respectively Y'') denotes the ordinary second derivative of X (respectively Y) as a function of one variable. Divide (5.1) by XY and we get

$$\frac{X''}{X} = -\frac{Y''}{Y}.$$

The left-hand side depends only on x, the right-hand side only on y and by equality of the two sides, both expressions must be constant, equal to λ. This gives the two equations $X'' = \lambda X$, $Y'' = -\lambda Y$.

There are now three cases:

- $\lambda = 0$: In this case $X'' = 0$ so $X = Ax + B$, $Y'' = 0$ so $Y = Cy + D$.
- $\lambda > 0$: In this case Y is a simple harmonic oscillator so

  $$Y = C\cos\left(y\sqrt{\lambda}\right) + D\sin\left(y\sqrt{\lambda}\right)$$

  and

  $$X = A\cosh\left(x\sqrt{\lambda}\right) + B\sinh\left(x\sqrt{\lambda}\right).$$

- $\lambda < 0$: In this case, X and Y are reversed (X becomes trigonometric, Y becomes hyperbolic):

$$Y = C\cosh\left(y\sqrt{-\lambda}\right) + D\sinh\left(y\sqrt{-\lambda}\right)$$

and

$$X = A\cos\left(x\sqrt{-\lambda}\right) + B\sin\left(x\sqrt{-\lambda}\right).$$

Boundary conditions

We have various possible boundary conditions we can apply along the edges of our rectangle.

- **Dirichlet boundary conditions:** where we fix the values of $\phi(x, 0)$, $\phi(x, b)$, $\phi(-a, y)$ or $\phi(a, y)$.
- **Neumann boundary conditions:** where we fix the values of $\frac{\partial\phi}{\partial x}$ along a vertical boundary or $\frac{\partial\phi}{\partial y}$ along a horizontal boundary. More generally, we would fix the directional derivative of ϕ in a direction normal to the boundary.

We can either have pure Dirichlet conditions, pure Neumann conditions or a mixture (called *mixed boundary conditions*).

Fitting solutions to boundary data

Now let's examine the various solutions we found by separation of variables and see what kinds of boundary condition we can satisfy. We will concentrate on Dirichlet. We denote by U the rectangular region $[0, a] \times [0, b]$ and by ∂U its boundary.

Fitting to corners

We impose

$$\phi(x, 0) = F_1(x) \qquad \phi(x, b) = F_2(x)$$
$$\phi(0, y) = F_3(y) \qquad \phi(a, y) = F_4(y)$$

Let's denote the corner values by:

$$S_{00} = F_1(0) = F_3(0) \qquad S_{10} = F_1(a) = F_4(0)$$
$$S_{01} = F_2(0) = F_3(b) \qquad S_{11} = F_2(a) = F_4(b)$$

and let's start by fitting the $\lambda = 0$ solutions

$$\phi_0(x, y) = (Ax + B)(Cy + D) = BD + ADx + BCy + ACxy$$

to this corner data. The four corner values give us four equations for A, B, C, D which we can solve and we get

$$\phi_0(x,y) = S_{00} + x\frac{(S_{10}-S_{00})}{a} + y\frac{(S_{01}-S_{00})}{b} + xy\frac{(S_{11}-S_{10}-S_{01}+S_{00})}{ab}$$

You can check that this satisfies the corner data. Now define

$$\theta(x,y) = \phi(x,y) - \phi_0(x,y)$$

and θ is a solution to Laplace's equation which solves some new boundary conditions, but crucially $\theta = 0$ at all the corners of U.

In other words, without loss of generality, our solution vanishes at the corners of U.

Patching solutions

We've seen that we can modify F_1, F_2, F_3 and F_4 to $F_1^{\text{new}}(x) = F_1(x) - \phi_0(x,0)$, $F_2^{\text{new}}(x) = F_2(x) - \phi(x,b)$, $F_3^{\text{new}}(y) = F_3(x) - \phi_0(0,y)$ and $F_4^{\text{new}}(y) = F_4(y) - \phi_0(a,y)$ which vanish at the corners. So let's split our problem into four pieces: let's look for solutions $\theta_1, \theta_2, \theta_3, \theta_4$ such that:

$$\begin{aligned} &\theta_1(x,0) = F_1^{\text{new}}(x) && \theta_1(x,b) = 0 \\ &\theta_1(0,y) = 0 && \theta_1(a,y) = 0 \end{aligned}$$

and

$$\begin{aligned} &\theta_2(x,0) = 0 && \theta_2(x,b) = F_2^{\text{new}}(x) \\ &\theta_2(0,y) = 0 && \theta_2(a,y) = 0 \end{aligned}$$

etc. Note that we can only do this because $F_1^{\text{new}}(0) = F_1^{\text{new}}(a) = 0$, etc. otherwise the boundary conditions are discontinuous at the corners.

If we can find θ_1, θ_2, etc. then $\theta_1 + \theta_2 + \theta_3 + \theta_4$ is still a solution to Laplace's equation and satifies all the boundary conditions we wanted originally. I'll show you how to solve for θ_1: the other solutions can all be obtained the same way, or by being clever with symmetries of the rectangle.

Solve for θ_1

θ_1 will be written as a (possibly infinite) linear combination of separated solutions. We want these separated solutions individually to satisfy $X(0) = X(a) = 0$ to ensure that $\theta_1(0,y) = 0 = \theta_1(a,y)$.

If $\lambda > 0$ then $X(x) = A\cosh(x\sqrt{\lambda}) + B\sinh(x\sqrt{\lambda})$. The condition $X(0) = 0$ means that $A = 0$ and the condition $X(L) = 0$ means that $B = 0$. So there are no separated solutions with $\lambda > 0$ that work.

If $\lambda < 0$ then $X(x) = A\cos(x\sqrt{-\lambda}) + B\sin(x\sqrt{-\lambda})$. The condition $X(0) = 0$ means that $A = 0$ and the condition $X(L) = 0$ means that $B\sin(a\sqrt{-\lambda}) = 0$ which means that if $B \neq 0$ we need $a\sqrt{-\lambda} = n\pi$ for some $n \in \mathbf{Z}$. Thus $\lambda = -n^2\pi^2/a^2$.

We therefore want a combination

$$\theta_1(x,y) = \sum_{n=1}^{\infty} \left(C_n \sinh(n\pi y/a) + D_n \cosh(n\pi y/a)\right) \sin(n\pi x/a)$$

which fits the final two boundary conditions

$$\theta_1(x,0) = F_1(x),\ \theta_1(x,b) = 0.$$

These give us

$$F_1(x) = \sum_{n=1}^{\infty} D_n \sin(n\pi x/a)$$

so that D_n are the coefficients in the Fourier half-range sine series of $F_1(x)$ and

$$0 = \sum_{n=1}^{\infty} (C_n \sinh(n\pi b/a) + D_n \cosh(n\pi b/a)) \sin(n\pi x/a)$$

which means

$$C_n \sinh(n\pi b/a) + D_n \cosh(n\pi b/a) = 0$$

or

$$C_n = -D_n \coth(n\pi b/a).$$

Example 72. *If $F_1(x) = \sin(x)$ on $[0, a = \pi] \times [0, b = 5\pi]$ then $D_1 = 1$ and $C_1 = \coth(5\pi)$ so the solution is*

$$\theta_1(x,y) = (-\coth(5\pi)\sinh y + \cosh y)\sin x.$$

Finding the other θ_i

You could of course plough through this procedure for each θ. For instance, θ_2 is particularly easy: we need to solve

$$0 = \sum_{n=1}^{\infty} D_n \sin(n\pi x/a),$$

whch gives $D_n = 0$, and

$$F_2^{\text{new}}(x) = \sum_{n=1}^{\infty} C_n \sinh(n\pi b/a) \sin(n\pi x/a)$$

so that C_n is $1/\sinh(n\pi b/a)$ times the nth Fourier coefficient of $F_2^{\text{new}}(x)$.

Solving for θ_3 and θ_4 means you need the $\lambda > 0$ solutions so that Y is trigonometric, but otherwise the solutions are found exactly the same way.

Sometimes, using a bit of geometric intuition (e.g. reflecting the rectangle) will allow you to guess the solution.

5.2.5 Some examples

Example 73. *Here is an example. On the square* $[0,\pi]^2$*, take* $F_1(x) = \sin(x)$*,* $F_2(x) = 0$*,* $F_3(y) = 0$*,* $F_4(x) = \sin(y)$*.*

We get $\theta_1(x,y) = (-\coth(\pi)\sinh y + \cosh y)\sin x$ *as in the previous example. However, we could rearrange this using hyperbolic trigonometric identities as* $\frac{\sinh(\pi-y)\sin x}{\sinh \pi}$*. If we had instead begun with* $F_2(x) = \sin(x)$*,* $F_1(x) = 0$*, the solution would have been* $\frac{\sinh y \sin x}{\sinh \pi}$*. Clearly these two solutions are related by* $y \mapsto \pi - y$*, which is a reflection in the horizontal line which cuts the rectangle in two.*

We also get $\theta_4(x,y) = \frac{\sinh x \sin y}{\sinh \pi}$*. So the full solution is*

$$\phi(x,y) = \frac{\sinh x \sin y}{\sinh \pi} + \frac{\sinh(\pi - y)\sin x}{\sinh \pi}.$$

Example 74. *On the square* $[0,\pi]^2$*, take* $F_1(x) = 0$*,* $F_2(x) = x^2$*,* $F_3(y) = 0$*,* $F_4(y) = y^2$*.*

We start by finding a $\lambda = 0$ *solution to fit the corner data. Consider the function* xy*. This takes on the right values at the corners of our square (i.e. the values* $0, 0, 0, \pi$*). If* ϕ *solves our problem then* $\theta(x,y) = \phi(x,y) - xy$ *solves Laplace's equation with boundary values*

$$F_1^{new}(x) = 0,\ F_2^{new}(x) = x^2 - \pi x,\ F_3^{new}(y) = 0,\ F_4^{new}(y) = y^2 - \pi y$$

(in particular, the boundary values are zero at the corners). The sinusoidal Fourier series of $x^2 - \pi x$ *on* $[0,\pi]$ *(i.e. the half-range series obtained by extending it to an odd function on* $[-\pi,\pi]$*) is*

$$-\frac{8}{\pi}\sum_{k=1}^{\infty}\frac{\sin((2k-1)x)}{(2k-1)^3}$$

and because the values in the corners are now zero we can apply our earlier solution

$$-\frac{8}{\pi}\sum_{k=1}^{\infty}\left(\frac{\sin((2k-1)x)\sinh((2k-1)y)}{(2k-1)^3\sinh((2k-1)\pi)} + \frac{\sin((2k-1)y)\sinh((2k-1)x)}{(2k-1)^3\sinh((2k-1)\pi)}\right)$$

To get ϕ rather than $\phi(x,y) - xy$ we need to add xy so the final answer is

$$xy - \frac{8}{\pi}\sum_{k=1}^{\infty}\left(\frac{\sin((2k-1)x)\sinh((2k-1)y)}{(2k-1)^3\sinh((2k-1)\pi)} + \frac{\sin((2k-1)y)\sinh((2k-1)x)}{(2k-1)^3\sinh((2k-1)\pi)}\right).$$

5.3 Eigenfunction expansions

This section is nonexaminable. It attempts to answer the question "Why is $\sin(n\pi x/L)$ turning up all over the place?".

Consider the vector space Y of functions $[0,\pi] \to \mathbf{R}$ and the subspace

$$X = \{F\colon [0,\pi] \to \mathbf{R} \ : \ F(0) = F(\pi) = 0\}$$

of functions on $[0,\pi]$ which vanish at 0 and π. We can define a linear operator

$$\frac{d^2}{dx^2}\colon X \to Y$$

which takes F to the second derivative of F. Note that $F''(0)$ and $F''(\pi)$ need not vanish! Nonetheless, we can ask for eigenvalues and eigenvectors (eigenfunctions) of d^2/dx^2, that is functions $F \in X$ such that

$$F'' = d^2F/dx^2 = \lambda F$$

for some $\lambda \in \mathbf{R}$. First notice that

$$\int_0^\pi \lambda F^2 dx = \int_0^\pi FF''dx = -\int_0^\pi F'F'dx \leq 0$$

by integrating by parts. This implies $\lambda \leq 0$. The equation $F'' = \lambda F$ has solutions $A\sin(x\sqrt{-\lambda}) + B\cos(x\sqrt{-\lambda})$ and these vanish at 0 and π if and only if $B = 0$ and $\sqrt{-\lambda} = n\pi$ for some $n \in \mathbf{Z}$.

Now consider the heat equation in three variables

$$\partial\phi/\partial t = \Delta\phi$$

where $\phi(t,x,y)$ is now a function of three variables and Δ is the Laplacian. Seeking separated solutions $\phi = T(t)M(x,y)$ gives

$$T'/T = \Delta(M)/M = \lambda$$

where λ is constant and hence $T = e^{\lambda t}$, $\Delta(M) = \lambda M$. So what are the eigenfunctions of Δ? Well we worked them out on Sheet 7. As usual there's a discrete set of them M_n corresponding to a discrete collection of eigenvalues λ_n. To replace the usual Fourier expansion of a function F in terms of sines, there is now an *eigenfunction expansion* of F as a sum of eigenfunctions of M_n:

$$F(x,y) = \sum_n A_n M_n(x,y)$$

and $\Delta(F) = \sum_n A_n \lambda_n M_n(x,y)$. Thus a solution to the heat equation with initial conditions $\phi(0,x,y) = M(x,y)$ is given by

$$\sum_n A_n e^{\lambda_n t} M_n(x,y)$$

and the nth mode of F decays in time with rate λ_n.

Returning to 1-dimensional problems, we might try to solve the equation

$$\phi_t = f(x)\phi_{xx} + g(x)\phi_x + h(x)\phi$$

by separating variables $\phi(x,t) = T(t)X(x)$. The result is

$$T'/T = f(x)X''/X + g(x)X'/X + h(x) = \lambda$$

where λ is constant. Solving problems like

$$f(x)X'' + g(x)X' + h(x)X = \lambda X$$

is the subject of Sturm-Liouville theory. You've already met an equation like this on Sheet 4, namely Legendre's equation

$$X''(1-x^2) - 2xX' = -\ell(\ell+1)X$$

where $\lambda = \ell(\ell+1)$ for $\ell \in \mathbf{Z}$ are the eigenvalues.

The Sturm-Liouville theory gives conditions on f, g, h which ensure the existence of an infinite discrete set of eigenvalues λ_n with eigenfunctions y_n. The key properties which let us mimic Fourier theory are:

$$\int_0^\pi y_n(x) y_m(x) dx = \delta_{mn}$$

(orthogonality) like the integrals of sine and cosine which let us compute Fourier series and *completeness*, which asserts the existence of an expansion

$$F(x) = \sum_n A_n y_n(x)$$

for any (reasonable, e.g. square-integrable) function. This orthogonality can be understood as an infinite-dimensional analogue of the following theorem from linear algebra:

Theorem 75. *Suppose A is a symmetric matrix. If v and w are eigenvectors of A associated to different eigenvalues $\lambda \neq \mu$ then $v \cdot w = 0$.*

Proof. Consider

$$\begin{aligned} v \cdot Aw &= (A^T v) \cdot w \\ &= (Av) \cdot w \end{aligned}$$

Since w is a μ-eigenvector of A and v is a λ-eigenvector of A this equation becomes

$$\mu(v \cdot w) = \lambda(v \cdot w)$$

If $\mu \neq \lambda$ this implies $v \cdot w = 0$. □

The analogue of the dot product $\cdot$ for the space of functions on $[0, \pi]$ is the integral

$$\text{“}f \cdot g\text{”} = \int_0^\pi f(x)g(x)dx$$

In function space, A is replaced by an appropriate operator like d^2/dx^2. The analogue of the equation

$$v \cdot Aw = (Av) \cdot w$$

is obtained by integrating by parts twice:

$$\int_0^\pi f(x)\frac{d^2g}{dx^2}(x)dx = -\int_0^\pi \frac{df}{dx}(x)\frac{dg}{dx}(x)dx = \int_0^\pi \frac{d^2f}{dx^2}(x)g(x)dx.$$

(Note that we need $f(0) = f(\pi) = g(0) = g(\pi) = 0$ or other suitable boundary conditions for this integration to work without picking up boundary terms).

A better name for “symmetric” in this context is “self-adjoint”: an operator has an *adjoint operator* A^T defined by

$$(A^T v) \cdot w = v \cdot Aw$$

and self-adjointness means that $A = A^T$.

5.4 Discontinuities

As a final example, I want to tackle the problem of what happens when your equation becomes discontinuous. For example, maybe you're interested in the heat equation on a rod $[0, 2\pi]$ whose thermal diffusivity changes dramatically halfway along and you model this by using the heat equation

$$\frac{\partial \phi}{\partial t} = \sigma \frac{\partial^2 \phi}{\partial x^2}$$

where σ is the discontinuous function

$$\sigma(x) = \begin{cases} \sigma_1 & \text{if } x \in [0, \pi] \\ \sigma_2 & \text{if } x \in [\pi, 2\pi] \end{cases}$$

Let's figure out what the separated solutions are when $\sigma_1 = 1$ and $\sigma_2 = 2$ (so the right-hand half of the rod conducts heat more quickly than the left-hand half).

We impose the boundary conditions $\phi(0,t) = \phi(2\pi,t) = 0$. Let's not impose an initial condition and just look for the separated solutions. We will also need extra "boundary conditions" at the discontinuity, namely we will require our separated solutions and their first derivatives to be continuous at π.

Let's write

$$\phi(x,t) = \begin{cases} \phi_1(x,t) = X_1(x)T_1(t) & \text{on } x \in [0,\pi] \\ \phi_2(x,t) = X_2(x)T_2(t) & \text{on } x \in [\pi, 2\pi] \end{cases}$$

On $[0,\pi]$ we get $T_1' = \lambda T_1$, $X_1'' = \lambda X_1$ as usual. On $[\pi, 2\pi]$ we get $T_2' = \lambda T_2$ and $X_2'' = 2\lambda X_2$. Therefore the separated solutions we seek are

$$X_1(x) = A_1 \sin(x\sqrt{-\lambda}) + B_1 \cos(x\sqrt{-\lambda}),\ X_2(x) = A_2 \sin(x\sqrt{-\lambda/2}) + B_2 \cos(x\sqrt{-\lambda/2})$$

The boundary condition $\phi(0,t) = 0$ becomes

$$X_1(0) = 0 \ \Rightarrow\ B_1 = 0$$

and the boundary condition $\phi(2\pi,t) = 0$ becomes

$$X_2(2\pi) = 0 \ \Rightarrow A_2 \sin(2\pi\sqrt{-\lambda/2}) + B_2 \cos(2\pi\sqrt{-\lambda/2}) = 0$$

We have four constants left A_1, A_2, B_2, λ and only one equation connecting them. We need to impose continuity of X and X' at $x = \pi$ to fix these extra constants.

We have

$$\begin{aligned} &X_1(\pi) = A_1 \sin(\pi\sqrt{-\lambda}) \\ &X_2(\pi) A_2 \sin(\pi\sqrt{-\lambda/2}) + B_2 \cos(\pi\sqrt{-\lambda/2}) \end{aligned}$$

so

$$X_1(\pi) = X_2(\pi)$$

implies

$$A_1 \sin(\pi\sqrt{-\lambda}) = A_2 \sin(\pi\sqrt{-\lambda/2}) + B_2 \cos(\pi\sqrt{-\lambda/2})$$

and

$$\begin{aligned} &X_1'(\pi) = \sqrt{-\lambda} A_1 \cos(\pi\sqrt{-\lambda}) \\ &X_2'(\pi)\sqrt{-\lambda/2}\left(A_2 \cos(\pi\sqrt{-\lambda/2}) - B_2 \sin(\pi\sqrt{-\lambda/2})\right) \end{aligned}$$

so

$$X_1'(\pi) = X_2'(\pi)$$

implies

$$\sqrt{-\lambda} A_1 \cos(\pi\sqrt{-\lambda}) = \sqrt{-\lambda/2}\left(A_2 \cos(\pi\sqrt{-\lambda/2}) - B_2 \sin(\pi\sqrt{-\lambda/2})\right)$$

We already know that

$$A_2 \sin(2\pi\sqrt{-\lambda/2}) = -B_2 \cos(2\pi\sqrt{-\lambda/2})$$

i.e. $B_2 = -A_2 \tan(2\pi\sqrt{-\lambda/2})$, so the two new equations give

$$A_1 \sin(\pi\sqrt{-\lambda}) = A_2 \left(\sin(\pi\sqrt{-\lambda/2}) - \tan(2\pi\sqrt{-\lambda/2})\cos(\pi\sqrt{-\lambda/2})\right)$$
$$A_1\sqrt{-\lambda}\cos(\pi\sqrt{-\lambda}) = \sqrt{-\lambda/2}A_2 \left(\cos(\pi\sqrt{-\lambda/2}) + \tan(2\pi\sqrt{-\lambda/2})\sin(\pi\sqrt{-\lambda/2})\right)$$

Dividing the first by the second equation gives

$$\frac{1}{\sqrt{-\lambda}}\tan(\pi\sqrt{-\lambda}) = \frac{\sin(\pi\sqrt{-\lambda/2}) - \tan(2\pi\sqrt{-\lambda/2})\cos(\pi\sqrt{-\lambda/2})}{\sqrt{-\lambda/2}\left(\cos(\pi\sqrt{-\lambda/2}) + \tan(2\pi\sqrt{-\lambda/2})\sin(\pi\sqrt{-\lambda/2})\right)}$$

so

$$\begin{aligned}\tan(\pi\sqrt{-\lambda}) &= \sqrt{2}\frac{\tan(\pi\sqrt{-\lambda/2}) - \tan(2\pi\sqrt{-\lambda/2})}{1 + \tan(2\pi\sqrt{-\lambda/2})\tan(\pi\sqrt{-\lambda/2})} \\ &= -\sqrt{2}\tan\left(\pi\sqrt{-\lambda/2}\right)\end{aligned}$$

So λ has to satisfy this bizarre equation! If we plot the graphs of $\tan(\pi\sqrt{-\lambda})$ and $-\sqrt{2}\tan\left(\pi\sqrt{-\lambda/2}\right)$ as functions of λ, see Figure 5.1, we see that they intersect at a discrete set of points with abscissa λ_n. Therefore we get separated solutions with these λ_n and the other constants can be calculated from the relations we have already found.

This ceases to be something we can do explicitly because there is no closed form expression for the numbers λ_n.

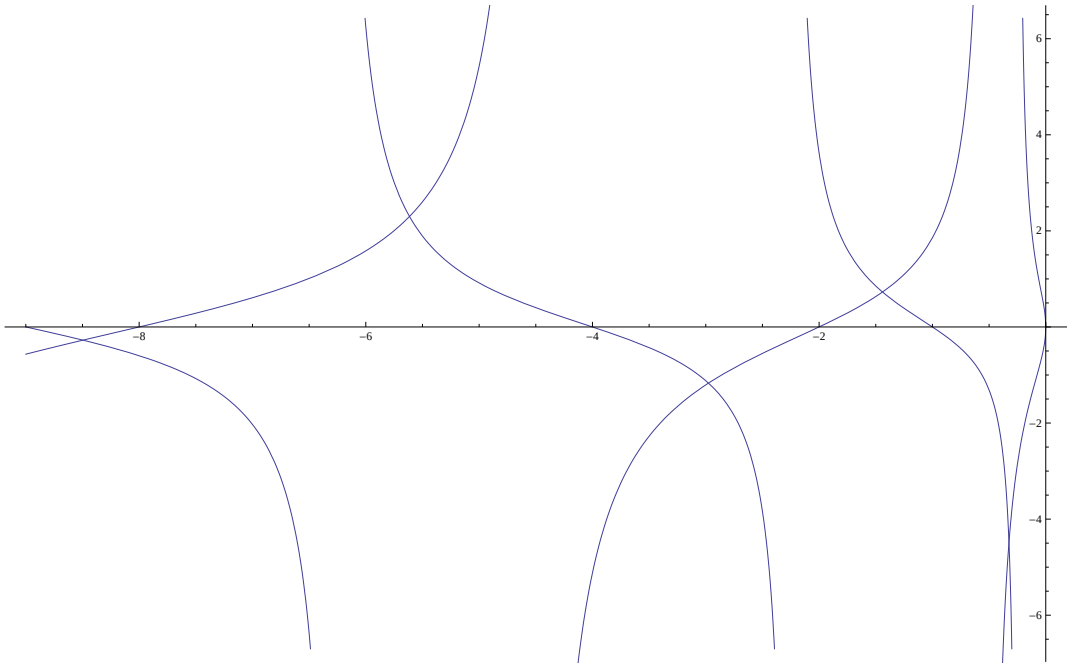


Figure 5.1: The graphs of $\tan(\pi\sqrt{-\lambda})$ and $-\sqrt{2}\tan\left(\pi\sqrt{-\lambda/2}\right)$ as functions of λ. The x-coordinates of the (discrete set of) intersection points are the numbers we are calling λ_n.

Chapter 6

The wave equation

6.1 Derivation

Let's consider a uniform one-dimensional string stretched out straight between two points $(0,0)$ and $(L,0)$. Make a small perturbation of the string so that it follows the graph of a function $u\colon [0,L] \to \mathbf{R}$ with $u(0) = u(L) = 0$. We will let go of the string and allow it to vibrate. At time t and above the point $x \in [0,L]$ the height of the string will be $\phi(x,t)$. Here's a claim which we won't justify, but which should be intuitively appealing because of Hooke's law which relates tension and lengthening:

Claim 76. *Assume that the perturbation ϕ is small and assume the same of its derivatives. The potential energy in this perturbed string is (to a good approximation) proportional the difference between its length and the length L of the unperturbed string, i.e. equal to*

$$\tau \int_0^L \left(\sqrt{1 + \left(\frac{\partial \phi}{\partial x} \right)^2} - 1 \right) dx$$

where τ is the tension per unit of lengthening. To lowest order in f (by Taylor expanding the integrand) this is approximately

$$\frac{\tau}{2} \int_0^L \left(\frac{d\phi}{dx} \right)^2 dx.$$

Moreover, the total kinetic energy of the string, once it is in motion, is to a good approximation

$$\frac{\rho}{2} \int_0^L \left(\frac{d\phi}{dt} \right)^2 dx$$

where ρ is the density in units of mass per unit length (constant by the assumption of uniformness).

We saw on Problem Sheet 5 that, in the Lagrangian formulation of mechanics, equations of motion can be derived by minimising the Lagrangian given by the difference between kinetic and potential energy. Using this principle, we will derive the equation of motion for the string by extremising the functional

$$\int_0^L \left(\frac{\rho}{2} \left(\frac{\partial \phi}{\partial t} \right)^2 - \frac{\tau}{2} \left(\frac{\partial \phi}{\partial x} \right)^2 \right) dx$$

The (two-variable) Euler-Lagrange equation for this functional is

$$\frac{\partial}{\partial t} \rho \left(\frac{\partial \phi}{\partial t} \right) - \frac{\partial}{\partial x} \tau \left(\frac{\partial \phi}{\partial x} \right)$$

or

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = \frac{\partial^2 \phi}{\partial x^2}$$

where $c = \sqrt{\tau/\rho}$. We will see that c has the interpretation of speed for the waves of vertical disturbance that start to move up and down the string.

6.2 Boundary conditions

We have already seen how to solve the wave equation (and more generally linear hyperbolic equations with constant coefficients) in Section 3.3.2. The general solution is

$$\phi(x,t) = F(x+ct) + G(x-ct)$$

for some arbitrary functions F and G. The only remaining task is to compute F and G from the initial data of the string at time $t = 0$. Since there are two arbitrary functions F and G (or because the wave equation has second order) we need to know more than just the initial profile u of the string. We need extra boundary or initial conditions. Above, we assumed fixed endpoints $\phi(0,t) = \phi(L,t) = 0$, but there are other possibilities:

- Maybe the string is infinitely long, but we know $v(x) \equiv \frac{\partial \phi}{\partial t}(x,0)$. If we were just letting go of the string then it would be stationary at the instant we let it go (though it would immediately start to accelerate). In other words we would have $v \equiv 0$. However, we allow ourselves more general functions v.

- Maybe the string extends infinitely far in only one direction (i.e. along the interval $[0,\infty)$, but is fixed so that $\phi(0,t) = 0$.

More generally there are all sorts of boundary conditions you could stick on your string. We'll just concern ourselves with the fixed endpoints and infinite string settings.

6.2.1 The bounded string: Fourier's solution

We deal first with the case of a string stretched between 0 and L and fixed at height zero. That is we impose the boundary conditions

$$\phi(0,t) = \phi(L,t) = 0.$$

We will seek solutions to the wave equation by separating variables, that is looking for solutions $\phi(x,t) = X(x)T(t)$. Substituting this into the wave equation gives:

$$X''T = \frac{1}{c^2} XT''$$

and so we only need to solve simple harmonic oscillator equations

$$X'' = \lambda X$$
$$T'' = c^2 \lambda T$$

Claim 77. *There are no nontrivial solutions satisfying the boundary conditions*

$$\phi(0,t) = \phi(L,t) = 0$$

if $\lambda \geq 0$.

Proof. When $\lambda = 0$ we have $X'' = 0$ which means $X = Ax + B$. For this to vanish at $x = 0, L$ we need $A = B = 0$. Similarly if $\lambda > 0$ then $X = A\cosh(x\sqrt{\lambda}) + B\sinh(x\sqrt{\lambda})$; setting $x = 0, L$ and requiring $\phi(0,t)$ and $\phi(L,t)$ to vanish implies

$$A\cosh(0) + B\sinh(0) = 0 \;\Rightarrow\; A = 0$$

and

$$B\sinh(L\sqrt{\lambda}) = 0 \;\Rightarrow B = 0.$$

Hence $A = B = 0$ and the solution is trivial. □

In the remaining case $\lambda < 0$ there are many solutions when $\sqrt{\lambda} = k\pi/L$ as usual

$$X(x) = \sin(k\pi x/L), \qquad T(t) = C_k \cos(kc\pi t/L) + D_k \sin(kc\pi t/L).$$

Because the equation is second-order in time, we need to specify both $f(x) = \phi(x,0)$ and $g(x) = \frac{\partial \phi}{\partial t}(x,0)$. Suppose that

$$f(x) = \sum_{k=1}^{\infty} F_k \sin(k\pi x/L)$$
$$g(x) = \sum_{k=1}^{\infty} G_k \sin(k\pi x/L)$$

are the half-range sinusoidal Fourier series for f and g. Then if

$$\phi(x,t) = \sum_{k=1}^{\infty} (C_k \cos(kc\pi t/L) + D_k \sin(kc\pi t/L)) \sin(n\pi x/L)$$

we see that $f(x) = \phi(x,0)$ implies

$$\sum_{k=1}^{\infty} F_k \sin(k\pi x/L) = \sum_{k=1}^{\infty} C_k \sin(n\pi x/L)$$

and $g(x) = \partial\phi/\partial t(x,0)$ implies

$$\sum_{k=1}^{\infty} G_k \sin(k\pi x/L) = -\sum_{k=1}^{\infty} \frac{k\pi c}{L} D_k \sin(n\pi x/L)$$

so $C_k = F_k$ and $D_k = LG_k/k\pi c$. The solution is therefore

$$\phi(x,t) = \sum_{k=1}^{\infty} \left(F_k \cos(kc\pi t/L) + \frac{LG_k}{k\pi c} \sin(kc\pi t/L) \right) \sin(k\pi x/L).$$

This has an obvious oscillatory behaviour in time which is characteristic of solutions to the wave equation, unlike the heat equation where oscillatory behaviour is supressed exponentially. Note that a Fourier mode with higher spatial frequency has correspondingly higher temporal frequency (this is because X and T satisfy simple harmonic motion with constants differing only by a factor of $1/c^2$).

The different summands are called the *normal modes of vibration*: $k = 1$ is called the fundamental mode, $k = 2$ is called the second harmonic, $k = 3$ the third harmonic, etc. Our claim about frequency means that a string vibrating in its kth mode will move up and down with frequency $k\pi c/L$.

Example 78. *Let's do the example of a plucked string, where the initial condition is* $f(x) = \begin{cases} mx & \text{if } x \le L/2 \\ mL/2 - mx & \text{if } x \ge L/2 \end{cases}$ *and* $g(x) = 0$. *The half-range Fourier sine series of* $f(x)$ *is*

$$\sum_{n=1}^{\infty} \frac{2mL}{n\pi} \left(\frac{2}{n\pi} \sin(n\pi/2) + \frac{\cos(n\pi) - \cos(n\pi/2)}{2} \right) \sin\left(\frac{n\pi x}{L}\right)$$

and the Fourier expansion of G *vanishes, so the solution is*

$$\phi(x,t) = \sum_{n=1}^{\infty} \frac{2mL}{n\pi} \left(\frac{2}{n\pi} \sin(n\pi/2) + \frac{\cos(n\pi) - \cos(n\pi/2)}{2} \right) \sin\left(\frac{n\pi x}{L}\right) \cos\left(\frac{n\pi ct}{L}\right).$$

6.2.2 The infinite string: D'Alembert's solution

In this section we will deal with the case of an infinite string, specifying $u(x) = \phi(x,0)$ and $v(x) = \frac{\partial \phi}{\partial t}(x,0)$. We have already seen how to solve the wave equation (and more generally linear hyperbolic equations with constant coefficients) in Section 3.3.2.

We know that

$$\phi(x,0) = u(x) = F(x) + G(x).$$

Since $\phi(x,t) = F(x+ct) + G(x-ct)$ we have

$$\begin{aligned} u(x) &= F(x) + G(x) \\ v(x) &= c(F'(x) - G'(x)) \end{aligned}$$

and so by the fundamental theorem of calculus

$$\int_0^x v(s)ds = c(F(x) - G(x)) + K$$

for some constant K, and hence

$$\begin{aligned} F(x) &= \frac{1}{2}\left(u(x) + \frac{1}{c}\left(\int_0^x v(s)ds - K\right)\right) \\ G(x) &= \frac{1}{2}\left(u(x) - \frac{1}{c}\left(\int_0^x v(s)ds - K\right)\right) \end{aligned}$$

which gives the general solution

$$\begin{aligned} \phi(x,t) &= F(x+ct) + G(x-ct) \\ &= \frac{1}{2}\left(u(x+ct) + u(x-ct)\right) + \frac{1}{2c}\left(\int_0^{x+ct} - \int_0^{x-ct}\right) v(s)ds \\ &= \frac{1}{2}\left(u(x+ct) + u(x-ct)\right) + \frac{1}{2c}\int_{x-ct}^{x+ct} v(s)ds \end{aligned}$$

due to D'Alembert.

6.2.3 Example

For a similar example worked out from first principles, see Section 3.3.2.

Example 79. *Suppose that* $\phi(x,0) = \sin x$ *and* $\partial\phi/\partial t(x,0) = \cos x$. *Then*

$$\begin{aligned} \phi(x,t) &= \frac{1}{2}(\sin(x+ct) + \sin(x-ct)) + \frac{1}{2c}\int_{x-ct}^{x+ct} \cos\xi d\xi \\ &= \frac{1}{2}(\sin(x+ct) + \sin(x-ct)) + \frac{1}{2c}(\sin(x+ct) - \sin(x-ct)) \\ &= \frac{1}{2}(1+1/c)\sin(x+ct) + \frac{1}{2}(1-1/c)\sin(x-ct). \end{aligned}$$

6.2.4 Propagating signals

Let's try to understand d'Alembert's solution a little better. Assume for the moment that $v \equiv 0$ so that the solution is

$$\phi(x,t) = \frac{1}{2}\left(u(x+ct) + u(x-ct)\right).$$

At each time t this is a superposition of two terms. Both terms have the same profile (shape), namely the profile of u (scaled down by a factor of two). As time progresses one of these profiles seems to move to the right ($u(x-ct)$, because its argument is decreasing as t increases) and the other ($u(x+ct)$) seems to move to the left. For instance, in Example 79 the crests of the $\frac{1+1/c}{2}\sin(x+ct)$ wave occur at

$$x + ct = 2\pi n + \pi/2$$

Let $x_n(t)$ denote the position of the nth crest at time t. This gives

$$x_n(t) = 2\pi n + \pi/2 - ct$$

so as t increases, $x_n(t)$ decreases and the crest moves to the left. We therefore call this a left-moving wave and the other a right-moving wave.

Example 80. *Suppose that*

$$u(x) = \begin{cases} 1 & \text{if } |x| < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Then two little square waves move off in either direction. Notice that an observer standing k metres to the left will only notice the arrival of the square wave after k/c seconds. In other words, signals propagate at a speed of precisely c.

We represent this causal relationship diagrammatically by drawing the (x,t)-plane (called a spacetime diagram). See Figure 6.1. In the pictures, $c = 1$ so that the slope of a line traced out by a wave front leaving the point $x = 0$ at time $t = 0$ is 1. We'll talk about the waves as light waves (just to aid the imagination). The lower part of the figure shows as dotted lines the light rays emitted forward in time from a point situated at the origin and the light which arrives at that point from the past. The *forward light cone* is the set of all points which could be reached from this point by travelling along light rays (maybe alternating between left- or right-moving rays by a cunning use of mirrors). On the other hand, light can never reach the regions $x > t$ because it would have to travel faster than light to get there. The *backward light-cone* is the set of all points from which light could conceivably reach the origin via a cunning system of mirrors.

In the upper part of the diagram you can see the support of the two square waves propagating to the left and right. The support of a function is the set of points where it's non-zero. The interval $[-1,1]$ at time 0 is the initial support of the square wave. As time progresses, the supports of the left- and right-moving square-waves move left and right along the dotted lines.

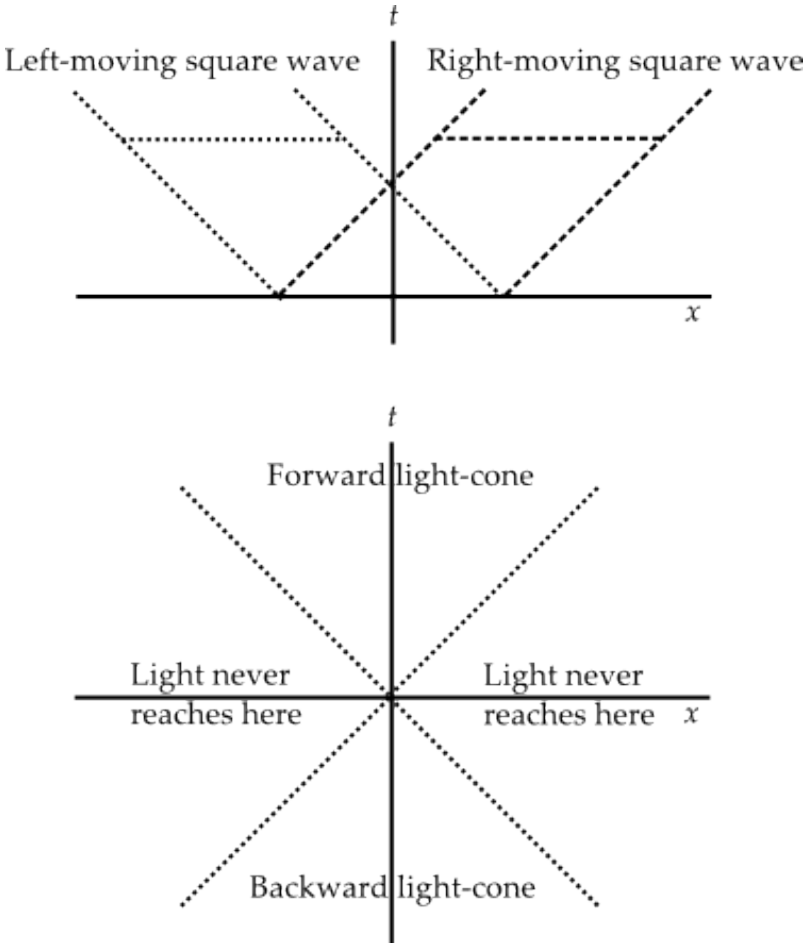


Figure 6.1: A spacetime diagram showing light rays (moving at speed $c = 1$) emitted from a point at the origin (lower part) and showing the motion of a left- and right-moving square wave (upper part).

6.2.5 Comparison of Fourier and D'Alembert solutions

We now have two ways of solving the wave equation: by separating variables à la Fourier and by a clever change of coordinates due to D'Alembert. It is easier to apply the separated variables method to boundary-value problems because it is not so easy to express the boundary conditions in terms of the arbitrary functions F and G from D'Alembert's solution. Of course the two solutions are related by a standard trigonometric identity:

$$\sin px \sin pct = \frac{\cos p(x - ct) - \cos p(x + ct)}{2}$$

i.e. a separated solution can be seen as a superposition of right- and left-moving waves.

6.3 Some simple waves on an infinite string

6.3.1 Simple waves on an infinite string

One particularly simple wave which we're allowed on an infinite string is the *simple harmonic wave*

$$\phi(x,t) = A\cos(k(x-ct)+B)$$

which we could re-express in terms of $\sin(k(x-ct))$ and $\cos(k(x-ct))$ if we were so inclined. The maximal height a is the *amplitude*, k is the *angular wavenumber*, $\lambda = 2\pi/k$ is the wavelength, $\omega = kc$ is the *angular frequency* and $T = 2\pi/\omega$ is the period. Changing the constant B is called shifting the *phase* of the wave. This wave is moving to the left (as time increases, $x-ct$ is decreasing) and you could create a right-moving wave by using $x+ct$.

A *standing wave*, which, as the name suggests, doesn't move, can be obtained by adding a left-moving and a right-moving simple harmonic wave with the same amplitude and frequency

$$\phi(x,t) = A\cos(k(x-ct)) + A\cos(k(x+ct)) = 2A\cos(kx)\cos(kct)$$

Another way to write the simple harmonic wave is as

$$\phi(x,t) = \mathrm{Re}\left(Ae^{iB}\exp(ik(x-ct))\right)$$

6.3.2 Reflection and transmission coefficients

Suppose that there are two strings tied together at the origin and extending off to infinity in either direction. Suppose that waves travel at speed c_- in the left-hand string and c_+ in the right-hand string. Imagine a simple harmonic wave incoming from the left

$$Ae^{ik(x-c_-t)}, \qquad x \le 0$$

and when it hits the discontinuity some of it is reflected and some of it is transmitted, resulting in a composition of three waves: incoming, reflected and transmitted, with amplitudes A, R, T and frequencies k, ℓ, m. In other words, the resulting wave has the form:

$$\phi(x,t) = \begin{cases} Ae^{ik(x-c_-t)} + R^{i\ell(x+c_-t)} & \text{when } x \le 0 \\ Te^{im(x-c_+t)} & \text{when } x \ge 0 \end{cases}$$

Note that the reflected term is right-moving! This is certainly a solution to the wave equation away from $x = 0$: it's a linear superposition of solutions and the equation is linear. At $x = 0$ it's not clear that the wave equation even makes sense because ϕ might not be twice differentiable there. However, in physics people often make such guesses, claiming that the equation itself is discontinuous at the origin and therefore it doesn't matter that the solution isn't twice differentiable there: we simply don't care what's going on. There's some physical process going on which we don't understand and which reflects and transmits the waves, all we can do is observe the reflected and transmitted waves as they travel out, away from the origin, to regions where we understand the wave equation! This is characteristic of *scattering calculations*.

Despite our ignorance, we will try to calculate R, T, ℓ and m in terms of A and k. To do this we need to impose some assumptions about ϕ at $x = 0$: namely, we want ϕ and its first x-derivative to be continuous at $x = 0$. Then

$$Ae^{-ikc_-t} + Re^{i\ell c_-t} = Te^{-imc_+t}$$

and

$$kAe^{ikc_-t} + \ell Re^{i\ell c_-t} = mTe^{-imc_+t}.$$

Since e^{int} are linearly independent in the space of all complex-valued functions we need $kc_- = -\ell c_- = mc_+$, that is

$$\ell = -k, \qquad m = kc_-/c_+.$$

The two equations become

$$A + R = T$$

and

$$k(A - R) = kTc_-/c_+$$

which we can solve to get

$$T = \frac{2Ac_+}{c_- + c_+}, \qquad R = \frac{A(c_+ - c_-)}{c_- + c_+}.$$

Part III

Appendices

Appendix A

Recap of Fourier theory

A.1 Fourier series

Let $L \in (0, \infty)$ be a positive real number. Fourier theory is concerned with functions $f\colon \mathbf{R} \to \mathbf{R}$ which are periodic with period $2L$ (in the sense that $f(x + 2L) = f(x)$) and the attempt to express them in the form

$$f(x) = c + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi x}{L}\right) + \sum_{n=1}^{\infty} b_n \cos\left(\frac{n\pi x}{L}\right). \tag{A.1}$$

The letters m and n will always stand for integers greater than or equal to one.

$$\int_{-L}^{L} \sin\left(\frac{m\pi x}{L}\right) \sin\left(\frac{n\pi x}{L}\right) dx = \begin{cases} 0 \text{ if } m \neq n, \\ L \text{ otherwise.} \end{cases} \tag{A.2}$$

$$\int_{-L}^{L} \cos\left(\frac{m\pi x}{L}\right) \cos\left(\frac{n\pi x}{L}\right) dx = \begin{cases} 0 \text{ if } m \neq n, \\ L \text{ otherwise.} \end{cases} \tag{A.3}$$

$$\int_{-L}^{L} \sin\left(\frac{m\pi x}{L}\right) \cos\left(\frac{n\pi x}{L}\right) dx = 0 \tag{A.4}$$

Assume for one moment that we are allowed to swap integrals and sums[1], for instance:

$$\int_{-L}^{L} \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi x}{L}\right) dx = \sum_{n=1}^{\infty} \int_{-L}^{L} a_n \cos\left(\frac{n\pi x}{L}\right) dx.$$

[1] You will see justification for this kind of operation next term in Analysis 4. Specifically, we need to assume that the Fourier series converges uniformly to f.

Then a function $f(x)$ with Fourier series given by (A.1) has Fourier coefficients

$$c = \frac{1}{2L}\int_{-L}^{L} f(x)dx \tag{A.5}$$

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos\left(\frac{n\pi x}{L}\right)dx \tag{A.6}$$

$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin\left(\frac{n\pi x}{L}\right)dx \tag{A.7}$$

So when do Fourier series exist? Providing f is Riemann-integrable over $[-L, L]$ you can make sense of the integrals (A.6) and (A.7) and write down a Fourier series (see Question 5 above).

But when do they actually have anything to do with the function you started with? Let's define the partial sums

$$S_N(x) = c + \sum_{n=1}^{N} a_n \cos\left(\frac{n\pi x}{L}\right) + \sum_{n=1}^{N} b_n \sin\left(\frac{n\pi x}{L}\right).$$

Theorem 81. *If $f\colon \mathbf{R} \to \mathbf{R}$ is a piecewise continuously differentiable function with period $2L$ then it admits a Fourier series* (A.1) *which converges pointwise to f, i.e. for all $x \in \mathbf{R}$ such that f is differentiable at x, the difference*

$$f(x) - S_N(x)$$

tends to zero as $N \to \infty$. Moreover, when x_0 is a point where f is discontinuous, the Fourier sums converge to

$$S_N(x_0) \to \lim_{\epsilon \to 0} \frac{f(x_0 - \epsilon) + f(x_0 + \epsilon)}{2}.$$

A.2 Square-integrable functions and Parseval's theorem

Theorem 82 (Parseval's theorem)**.** *If $f\colon \mathbf{R} \to \mathbf{R}$ is periodic with period $2L$ and square-integrable over $[-L, L]$ (meaning that the integral*

$$\int_{-L}^{L} |f(x)|^2 dx$$

exists and is finite) then

$$S_N \to f$$

in the very weak sense that

$$\int_{-L}^{L} |f(x) - S_N(x)|^2 dx \to 0$$

as $N \to \infty$. In terms familiar from statistics, you can think of S_N as a least-squares approximation to f (but by Fourier sums, rather than polynomials). Consequently,

$$2c^2 + \sum_{n=1}^{\infty}(a_n^2 + b_n^2) = \frac{1}{L}\int_{-L}^{L} |f(x)|^2 dx. \tag{A.8}$$

A.3 Interpretation

The way you should think about what's going on is by analogy with finite-dimensional vector spaces. With a finite-dimensional vector space V you can pick an inner product

$$\langle , \rangle \colon V \times V \to \mathbf{R}$$

which, given two vectors, outputs a number. The vectors v, w are *orthogonal* if $\langle v, w\rangle = 0$. A basis $e_1, \ldots, e_n$ is orthonormal if $\langle e_i, e_j\rangle = \delta_{ij}$.

In Fourier theory the relevant vector space is the space of periodic functions with period $2L$, square-integrable on $[-L, L]$. The inner product of f and g is

$$\frac{1}{L}\int_{-L}^{L} f(x)g(x)dx.$$

Note that this is a well-defined integral: if the integrals $\int_{-L}^{L} f(x)^2 dx$ and $\int_{-L}^{L} g(x)^2 dx$ are both finite then by the Cauchy-Schwarz inequality,

$$\left|\int_{-L}^{L} f(x)g(x)dx\right|^2 \leq \left(\int_{-L}^{L} f(x)^2 dx\right)\left(\int_{-L}^{L} g(x)^2 dx\right),$$

the inner product is also finite.

With respect to this inner product the "vectors" (i.e. functions)

$$\frac{1}{\sqrt{2L}}, \qquad \frac{1}{\sqrt{L}}\cos\left(\frac{m\pi x}{L}\right), \qquad \frac{1}{\sqrt{L}}\sin\left(\frac{m\pi x}{L}\right), \qquad m \geq 1$$

are all orthonormal. However, they do not form a basis! Being a basis for a vector space means that any other vector (in this case function) can be written as a linear combination of a **finite** number of basis elements. The Fourier coefficients are the orthogonal projections of a given square-integrable function f onto the basis directions and we have certainly seen functions with infinite Fourier series.

The right way to think about this is that the space of Fourier sums (i.e. Fourier series with only finitely many nonzero coefficients) is **dense** in the space of all square-integrable functions, and that you can approximate (in the least-squares sense!) arbitrary square-integrable functions by infinite sequences of Fourier sums S_N and let $N \to \infty$.

Appendix B

Sage code for diagrams

B.1 Figure 1.1

```
y=var('y')
x=var('x')
t=var('t')
P=plot3d(x*y/(x^2+y^2),(x,-1,1),(y,-1,1),plot_points=[150,150],frame=False)
Ax=parametric_plot(vector((t,0,0)),(t,-1.1,1.1), thickness=3, color='black')
Ay=parametric_plot(vector((0,t,0)),(t,-1.1,1.1), thickness=3, color='black')
Az=parametric_plot(vector((0,0,t)),(t,-0.7,0.7), thickness=3, color='black')
show(P+Ax+Ay+Az)
```

B.2 Figure 1.2

```
x=var('x')
y=var('y')
t=var('t')
P=plot3d(x^3-x-y^2,(x,-1,1),(y,-1,1),opacity=0.8,plot_points=[20,20],frame=False)
Q=plot3d(-x,(x,-0.4,0.4),(y,-0.4,0.4),opacity=0.8,color='gray',
    plot_points=[10,10],frame=False)
R=parametric_plot3d((0,0,0),(t,0,1),color='black',thickness=15)
show(P+Q+R)
```

B.3 Figure 1.3

```
y=var('y')
x=var('x')
t=var('t')
P=plot3d(x^4+y^4-x^2-y^2,(x,-0.5,1),(y,-0.5,1),plot_points=[20,20],
    opacity=0.85,frame=False)
Q=plot3d(0,(x,-0.3,0.3),(y,-0.3,0.3),color='gray',plot_points=[20,20],
    opacity=0.8,frame=False)
Ax=parametric_plot(vector((t,0,0)),(t,-1.1,1.1), thickness=3, color='black')
Ay=parametric_plot(vector((0,t,0)),(t,-1.1,1.1), thickness=3, color='black')
Az=parametric_plot(vector((0,0,t)),(t,-0.7,0.7), thickness=3, color='black')
R=parametric_plot3d((0,0,0),(t,0,1),color='black',thickness=15)
show(P+Q+R)
```

B.4 Figure 1.4

```
x=var('x')
y=var('y')
P=plot3d(x^2-y^2,(x,-1,1),(y,-1,1),opacity=0.8,plot_points=[20,20],
    frame=False)
Ax=parametric_plot(vector((t,0,0)),(t,-1.1,1.1), thickness=3,
    color='black')
Ay=parametric_plot(vector((0,t,0)),(t,-1.1,1.1), thickness=3,
    color='black')
Az=parametric_plot(vector((0,0,t)),(t,-0.7,0.7), thickness=3,
    color='black')
show(P+Ax+Ay+Az)
```

B.5 Figures 4.1 and 4.2

B.5.1 Parts (a)

Use one or other of the `show` commands.

```
x=var('x')
y=var('y')
z=var('z')
t=var('t')
u=var('u')
p=plot_vector_field3d((1,z,0),(x,-1,1),(y,-1,1),(z,-1,1),frame=False)
q=plot_vector_field3d((1,z,0),(x,-1,1),(y,-1,1),(z,0,1),frame=False)
r=parametric_plot3d((t,u+t*u,u),(t,0,1),(u,-1,1),plot_points=[10,10],
    opacity=0.8,frame=False)
s=parametric_plot3d((t,u+t*u^2,u^2),(t,0,1),(u,-1,1),plot_points=[10,10],
    opacity=0.8,frame=False)
Ax=parametric_plot(vector((t,0,0)),(t,-1.1,1.1),
    thickness=3, color='black')
Ay=parametric_plot(vector((0,t,0)),(t,-1.1,1.1),
    thickness=3, color='black')
Az=parametric_plot(vector((0,0,t)),(t,-0.7,0.7),
    thickness=3, color='black')
Az2=parametric_plot(vector((0,0,t)),(t,0,0.7),
    thickness=3, color='black')
show(p+r+Ax+Ay+Az,mesh=True)
show(q+s+Ax+Ay+Az2,mesh=True)
```

B.5.2 Figure 4.1(b)

```
t=var('t')
p1=parametric_plot((t,0),(t,-1.2,1),axes=False)
p2=parametric_plot((t,0.1+0.1*t),(t,-1.2,1),axes=False)
p3=parametric_plot((t,0.2+0.2*t),(t,-1.2,1),axes=False)
p4=parametric_plot((t,0.3+0.3*t),(t,-1.2,1),axes=False)
p5=parametric_plot((t,0.4+0.4*t),(t,-1.2,1),axes=False)
p6=parametric_plot((t,0.5+0.5*t),(t,-1.2,1),axes=False)
p7=parametric_plot((t,0.6+0.6*t),(t,-1.2,1),axes=False)
p8=parametric_plot((t,0.7+0.7*t),(t,-1.2,1),axes=False)
p9=parametric_plot((t,0.8+0.8*t),(t,-1.2,1),axes=False)
pA=parametric_plot((t,0.9+0.9*t),(t,-1.2,1),axes=False)
pB=parametric_plot((t,1+1^2*t),(t,-1.2,1),axes=False)
q1=parametric_plot((t,0),(t,-1.2,1),axes=False)
q2=parametric_plot((t,-0.1-0.1*t),(t,-1.2,1),axes=False)
q3=parametric_plot((t,-0.2-0.2*t),(t,-1.2,1),axes=False)
q4=parametric_plot((t,-0.3-0.3*t),(t,-1.2,1),axes=False)
q5=parametric_plot((t,-0.4-0.4*t),(t,-1.2,1),axes=False)
q6=parametric_plot((t,-0.5-0.5*t),(t,-1.2,1),axes=False)
```

```
q7=parametric_plot((t,-0.6-0.6*t),(t,-1.2,1),axes=False)
q8=parametric_plot((t,-0.7-0.7*t),(t,-1.2,1),axes=False)
q9=parametric_plot((t,-0.8-0.8*t),(t,-1.2,1),axes=False)
qA=parametric_plot((t,-0.9-0.9*t),(t,-1.2,1),axes=False)
qB=parametric_plot((t,-1-1*t),(t,-1.2,1),axes=False)
Ax=parametric_plot(vector((t,0)),(t,-1.1,1.1),
    thickness=3, color='black')
Ay=parametric_plot(vector((0,t)),(t,-1.1,1.1),
    thickness=3, color='black')
show(p1+p2+p3+p4+p5+p6+p7+p8+p9+pA+pB+q1+q2+q3+q4+q5+q6+
    q7+q8+q9+qA+qB+Ax+Ay)
```

B.5.3 Figure 4.2(b)

```
t=var('t')
p1=parametric_plot((t,0),(t,-1.2,1),axes=False)
p2=parametric_plot((t,0.1+0.1^2*t),(t,-1.2,1),axes=False)
p3=parametric_plot((t,0.2+0.2^2*t),(t,-1.2,1),axes=False)
p4=parametric_plot((t,0.3+0.3^2*t),(t,-1.2,1),axes=False)
p5=parametric_plot((t,0.4+0.4^2*t),(t,-1.2,1),axes=False)
p6=parametric_plot((t,0.5+0.5^2*t),(t,-1.2,1),axes=False)
p7=parametric_plot((t,0.6+0.6^2*t),(t,-1.2,1),axes=False)
p8=parametric_plot((t,0.7+0.7^2*t),(t,-1.2,1),axes=False)
p9=parametric_plot((t,0.8+0.8^2*t),(t,-1.2,1),axes=False)
pA=parametric_plot((t,0.9+0.9^2*t),(t,-1.2,1),axes=False)
pB=parametric_plot((t,1+1^2*t),(t,-1.2,1),axes=False)
q1=parametric_plot((t,0),(t,-1.2,1),axes=False)
q2=parametric_plot((t,-0.1+0.1^2*t),(t,-1.2,1),axes=False)
q3=parametric_plot((t,-0.2+0.2^2*t),(t,-1.2,1),axes=False)
q4=parametric_plot((t,-0.3+0.3^2*t),(t,-1.2,1),axes=False)
q5=parametric_plot((t,-0.4+0.4^2*t),(t,-1.2,1),axes=False)
q6=parametric_plot((t,-0.5+0.5^2*t),(t,-1.2,1),axes=False)
q7=parametric_plot((t,-0.6+0.6^2*t),(t,-1.2,1),axes=False)
q8=parametric_plot((t,-0.7+0.7^2*t),(t,-1.2,1),axes=False)
q9=parametric_plot((t,-0.8+0.8^2*t),(t,-1.2,1),axes=False)
qA=parametric_plot((t,-0.9+0.9^2*t),(t,-1.2,1),axes=False)
qB=parametric_plot((t,-1+1^2*t),(t,-1.2,1),axes=False)
Ax=parametric_plot(vector((t,0)),(t,-1.1,1.1),
    thickness=3, color='black')
Ay=parametric_plot(vector((0,t)),(t,-1.1,1.1),
    thickness=3, color='black')
r=implicit_plot(1+4*x*y==0,(x,-1,1),(y,-1,1),color='red',
    frame=False)
show(p1+p2+p3+p4+p5+p6+p7+p8+p9+pA+pB+q1+q2+q3+q4+q5+q6+
    q7+q8+q9+qA+qB+Ax+Ay+r)
```

B.6 Figure 4.4

```
t=var('t')
p_1=parametric_plot((2*cos(t),sin(t)),(t,0,2*pi))
p_2=parametric_plot((2*cos(t)-0.2*cos(t)/(1+sin(t)*sin(t)),sin(t)-0.2*2*sin(t)/(1+sin(t)*sin(t))),(t,0,2*pi))
p_3=parametric_plot((2*cos(t)-0.5*cos(t)/(1+sin(t)*sin(t)),sin(t)-0.5*2*sin(t)/(1+sin(t)*sin(t))),(t,0,2*pi))
p_4=parametric_plot((2*cos(t)-0.7*cos(t)/(1+sin(t)*sin(t)),sin(t)-0.7*2*sin(t)/(1+sin(t)*sin(t))),(t,0,2*pi))
p_5=parametric_plot((2*cos(t)-1*cos(t)/(1+sin(t)*sin(t)),sin(t)-1*2*sin(t)/(1+sin(t)*sin(t))),(t,0,2*pi))
p_6=parametric_plot((2*cos(t)-1.2*cos(t)/(1+sin(t)*sin(t)),sin(t)-1.2*2*sin(t)/(1+sin(t)*sin(t))),(t,0,2*pi))
show(p_1+p_2+p_3+p_4+p_5+p_6)
```